Rats Hunt Stag

Integration of social and economic information drives cooperation in a collective decision making task.

Scott Rennie

Dissertation presented to obtain the Ph.D degree in neurosceince

Instituto de Tecnologia Química e Biológica | Universidade Nova de Lisboa

Research work coordinated by:

Oeiras,
August, 2016

INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
/UNL

Knowledge Creation

Rats Hunt Stag

The integration of social and economic information drives cooperation in a collective decision making task.

Scott Mckenzie Rennie

A Dissertation
Presented to the faculty
Of Universidade Nova de Lisboa
In Candidacy for the degree
Of Doctor in Neuroscience

Under the supervision of:
Dr Marta A Moita

August 2016

To Pooja

Acknowledgements

(I admit, I don't quite know how, and I have tried, to find the words that rise to task express the depths of gratitude to those that have given their time, and love and support over these last eight years. Explaining this is more daunting a task than all the data contained within these pages. I want to express at the very least that they did not just help, but were necessary in this process. To each of them I will express in person my thanks, and my apologies to those I have failed to list here. I will mention in detail those that collaborated directly on this project)

Marta. Who more than my supervisor, is my friend, and continues to show me, daily, how to be not only a better scientist, but a better person.

Eric: Who is only in these acknowledgements because convention prevents me from making him co-author of this thesis.

Alexandra: For collecting the all cooperate stooge, and decoupled condition data and for sharing the frustrations of these experiments, and without doubt, of working with me.

To my lab, all of whom helped in their unique way, particularly Cristina Marquez for her collaboration and attention to detail.

To my thesis committee Zach and Eugenia, and also Goncalo Polavieja, and Alfonso for their guidance, Adam also, I think.

Susana. Were it not for your strength and friendship this thesis would never have been completed.

Zach, for unwavering inspiration, friendship, and for being the right DJ at the right time so frequently.

To Andreia for the support and friendship when it was most needed.

To my family for tolerating my absence all these years and supporting me nonetheless.

Sam: for lessons in how to get things done that I haven't got around to learning (muurgh)

Daniel: For his genius whether practical, nourishing, or both.

Tom: For many useful discussions and perhaps as many arguments.

Marina: For friendship, essential care and for important introductions.

To Pooja, whom I love.

Sumário

Tomadas de decisão social apresentam, discutívelmente, o problema mais complexo que um animal pode encontrar. Tomadas de decisão económicas colectivas requerem a integração de predições baseada em resultados de interações prévias, conjuntamente com predições geradas por informações sociais correntes. Muitas decisões económicas são feitas enquanto indivíduos interagem, contudo como a maneira com que animais percebem e expressam informação social afeta decisões económicas permanence amplamente ignorado. Portanto, desenvolvemos uma tarefa de dilema social, tradicionalmente focada em como resultados de experiências prévias afectam escolhas, permitindo o acesso de cada rato jogador à informação social proximal. Afim de explorar como pares de ratos integram tais formas de informação, desenvolvemos um ensaio baseado em um labirinto em T duplo para testar jogos de dilema social 2x2 e estabelecemos uma tarefa de escolha social que corresponde à um jogo de Caça ao Veado (Stag Hunt, aqui SH) de alto risco, onde cada animal possui acesso completo à informação social enquanto fazem escolhas e recebem recompensas. No jogo SH existem dois equilíbrios de Nash: cooperação mútua, produzindo a recompensa mais alta com o maior risco (a recompensa depende da reciprocação do outro) e deserção mútua, que confere uma recompensa intermediária constante (independente do outro). Isso faz com que a escolha óptima para cada animal seja, em cada tentativa, fazer a escolha que acreditam que o oponente também fará.

Nós demonstramos que animais têm a capacidade de associar a sua própria escolha com àquela de animais fantoche controlados

experimentalmente, aumentando a sua recompensa de forma consistente com os termos económicos do SH. A seguir examinamos o comportamento livre de duplas de ratos na ausência de incentivos económicos e observamos uma tendência moderada à coordenação de suas escolhas e uma marcante preferência de alternar escolhas a cada tentativa.

A seguir, examinamos o comportamento de duplas de ratos, onde cada rato possui a opcão de escolher primeiro e desertar ou arriscar cooperar, ou então escolher em segundo e coordenar, realizando a mesma escolha que o oponente ou anti-coordenar, escolhendo diferentemente. Os ratos apresentaram uma habilidade robusta de coordenar em alternância desde a primeira sessão (65%) e gradualmente aumentar a tendência à enviesar mútuamente a alternância rumo à cooperação mútua (50% até a sessão 6, com probabilidade de 25%) significativamente ultrapassando a deserção mútua. Esse alto nível de cooperação mútua é suportado e caracterizado por altos níveis de tolerância à resultados baixos recebidos por cooperação não reciprocada. Assim que a cooperação estável foi estabelecida em todas as duplas nós removemos toda a informação social que resultou na maioria dos pares revertendo à estratégia dominada por deserção. Além disso, ao manipular a matriz de recompensas descobrimos que somente o histórico de recompensa, quandos as recompensas são independentes de outros comportamentos, não levam à preferência pela cooperação unilateral ou mutual. Isso sugere que a cooperação é baseada na capacidade dos animais prever resultados em cada tentativa como uma função da ação colaborativa colectiva. Nossos resultados

mostram que ratos aprendem a integrar a informação social corrente derivando de sua interação para cooperar em um jogo de coordenação de alto risco, para o qual eles usam ambas informações económicas e sociais. Acreditamos que este estudo pavimenta o caminho para os mecanismos subjacentes às decisões coletivas entre roedores em interação.

Abstract

Social decision-making presents arguably the most complex problem an animal can face. Collective, economic decision-making requires the integration of predictions based on the outcomes of prior interactions alongside predictions generated from ongoing social information. Many economic decisions are made as individuals interact with each other, however how the manner in which animals perceive and display social information affects economic decisions remains largely overlooked. Hence we developed a social dilemma task, traditionally focused on how experienced outcomes affect choices, but allow each rat player access to proximate social information. To explore how pairs of rats integrate these forms of information we developed a double T-maze assay for testing 2x2 social dilemma games and established a social choice task that corresponds to a high risk Stag Hunt (SH) game, where each animal has complete access to social information as they make choices and received rewards. In the SH game there are two Nash equilbria: mutual cooperation, yielding the highest reward at the greatest risk (reward depends on reciprocation by the other) and mutual defection, which provides a constant (independent of the other) intermediate reward. This makes the optimal choice for each animal, in each trial to make the choice they believe their opponent will also make.

We established that animals had the capacity to associate their own choice with that of experimentally controlled stooge animals to increase their reward in a manner consistent with the economic terms of the Stag Hunt. We then examined behaviour of freely behaving rat

dyads in the absence of economic incentives and observed a moderate tendency to coordinate their choices and a striking preference to alternate trial-by-trial choices.

Next we examined the behaviour of rat dyads, where each rat has the option to choose first and defect or risk cooperating, or choose second and coordinate, by making the same choice as their opponent or anti-coordinate, by choosing differently. Rats displayed a robust ability to coordinate in alternation from the first session (65%) and gradually increased their tendency to mutually bias their alternation towards mutual cooperation (50% by session 6, chance levels corresponds to 25%) significantly surpassing mutual defection. This high level of mutual cooperation was supported and characterized by high levels of tolerance to low outcomes received by unreciprocated cooperation. Once stable cooperation was established across all dyads we removed all social information, which resulted in the majority of pairs reverting to a strategy dominated by defection. Furthermore, by manipulating the payoff matrix we found that reward history alone, when rewards are independent of the other behaviors, does not lead to a preference for unilateral or mutual cooperation. This suggested that cooperation is based on animals being able to predict trial-by-trial outcomes as function of collaborative collective action. Our results show that rats learn to integrate the ongoing social information derived from their interaction to cooperate in a coordination high-risk game, for which they use both economic and social information. We believe this study paves the way for mechanistic underpinnings of collective decisions between interacting rodents.

TABLE OF CONTENTS

General Introduction

‘There’s a crack in everything’

Anthem, Leonard Cohen.

Social decision-making presents arguably the most complex problem an animal can face[1]. An animal's fate and fitness hinges upon its capacity to efficiently navigate the reciprocally intertwined probabilistic elements of its own perceptions, behaviour and that of the world around it. The study of social behavior is predicated on the position that the relationship between interacting conspecifics are in some manner distinct from that an animal has with the rest of its environment. At the heart of this distinction is the assertion that there is no other aspect of an animal's environment that is as likely, in any given moment, to share such a similar nexus of inclinations, capacities, proclivities, sensitivities and responses as a conspecific. The dilemma of how interacting conspecifics best manipulate their states and behaviours when both are engaging in much the same process is, in the context of such deep similarity, in principle, though perhaps in principle only, an unboundedly complex one.

A central question within the field of social behaviour and decision-making is that of cooperation. Cooperation is a mystery that refuses solution as much as it does definition. It can be employed in proximate or ultimate[2] terms, as a collective action[3] or individual behaviour, as being other regarding, or rationally self interested[4], and in the context of game theory it can be defined purely operationally[5]. Such diversity of definition has given rise to a cornucopia of approaches and experimental procedures of such variety that it is surprising that they share a collective interest[6].

The common intuitive consequence underlying this flea market of particulars, that individuals choose to interact in a manner where things work out collectively rather well, is (perhaps ironically) a defining aspect of human social behavior. This is particularly so when observed outside the constraints of kin, which presents the most commonly invoked factor supporting cooperation in other species[2,7,8]. In this sense humans represent an 'ideal model organism' for cooperation just as pigeons do for navigation[9], rats for olfaction[10–12] or starfish for regeneration[13]. The problem comes when we use the particular capacities of such an organism to etch out the very definition of that capacity and then set off to look for them in other species. While we have not done this for navigation, olfaction, or regeneration, it appears that we have for cooperation and indeed for social behavior in general. In so doing anthropomorphism and anthropocentrism cast a shadow upon the whole enterprise from the outset. The risk this brings is that we permit our conception of our own folk psychology to generate categories of social behaviour that are not appropriate to other species, or having done so, misinterpret the behaviour of those species as being supported by cognitive mechanisms much more aligned with our own folk psychology than they really are[14]. Exacerbating these problems is the very real concern that we may be mislead by confabulation and ascribe complex cognitive strategies and motivations to our own cooperative behaviours when more ergonomic explanations would suffice.

One way to mitigate the risks of such explanatory pitfalls is to, as best as one can, define a minimum set of conditions that circumscribe a tangible concept of cooperation, and where possible bring to bare

some quantifiable framework, whilst being careful not to drain the concept of meaning or application. The most widely and, arguably, successful approaches to cooperation by both experimental psychologists and evolutionary biologists have taken an economic perspective[2,7,15]. From this perspective the definition of cooperation that carries the least assumptions has two key components. Firstly it requires that the outcomes of individual action are dependent not only upon the actor's choice but are also, at least in part, dependent upon the actions of others. And secondly, that such actions lead to a more equitable distribution of available outcomes amongst actors[6,15]

Part of the confusion that often obscures what is meant by cooperation is that while psychological and evolutionary approaches are both trying, in their own way, to fully understand cooperative behaviour they emphasize different questions in their attempt to do so. The evolutionary approach to cooperation attempts to determine how behaviour that is beneficial to another would arise in a population[2,16]. It asks *why* do individuals cooperate. A psychological approach on the other hand asks *how* do individuals go about the act of cooperation. Of course to approach a complete understanding of behaviour both the how and they why must be addressed. Both proximate and ultimate explanations must be obtained[16,17]. This is particularly important when it comes to what is meant by 'outcomes' in the definition of cooperation above. From an evolutionary perspective outcomes are thought of in the currency of inclusive fitness, the number of offspring an individual produces that survive to reproductive age, or proxies thereof, whereas the psychological perspective considers such outcomes represent proximate rewards

such as food, or sex or safety. A complication here is that such proximate rewards are potentially meaningful proxies of fitness. This distinction between what is meant by outcomes gets right to the nub of the confusion. While these questions may be distinct, their role in a given behaviour or cognition is not. Natural selection leads animals to behave in the manner that they do, yet the manner in which animals behave is likely aimed at improving the proximate rather than optimization of the ultimate. For instance, the statement "females wish to mate with the highest status male the can find" can be made by both an evolutionary biologist and an experimental psychologist. Yet the former refers to behavioural strategies that maximize fitness and the latter to proximate psychological motivations[16]. To quickly address any potential confusion, this thesis will primarily follow a more traditional experimental psychological rather than evolutionary perspective foveating upon the proximate while keeping an eye upon the wider evolutionary context.

Cooperation thus defined; as an individual act that leads to collective benefit, has been then further categorized by the consequences too, and thusly considerations of, the actor. One such category is referred to as a mutualism. This refers to cooperative acts that incur no net costs and also provide immediate benefits, for example when two individuals groom each other at the same time.[18] This form of cooperation is quite common in diverse animal societies, particularly in cooperative hunting or breeding contexts[187]. Cooperative hunting can be defined as mutualism when the per capita rate of benefit of food intake within a particular hunting group exceeds that available to a solitary hunter[19]. In this example there is no reason for an animal

not to cooperate by joining the hunting group, given the opportunity to do so, as it gains more than it would lose, and does not incur a cost in the process. From the perspective of learning theory cooperation of this form is reinforcing as a cooperator is guaranteed benefit, and therefore cooperative acts are reinforced, increasing their likelihood of being repeated[20].

The absence of some cost to cooperation is what distinguishes mutualism from reciprocal altruism. Reciprocal altruism, or simply reciprocity, describes cases where an individual incurs a short-term cost for a cooperative act, but receives benefit in the long term. If we take the example of cooperative hunting and simply add the possibility that animals that engage in the hunt increase personal danger in the process or having expended energy engaging in a successful hunt, then run the risk of being excluded from the spoils, then this describes a type of cooperation that is much more complex in character.

In contrast to mutualisms, this more complex form of cooperation referred to as reciprocity has been much less frequently observed in animal societies[7,21]. This may be in part becasue it is not always clear precisely what are the outcomes and costs that may be being optimized by interactors which, in wild populations, complicates the identification of such cooperation. A further explanation for the apparent scarcity of this form of cooperation in animal societies is that it places greater cognitive demands upon the cooperator. Simply put, because now cooperation carries a cost that must be calculated. While there have been multiple examples in laboratory and natural populations such as tree swallows[22]. stickelbacks[23], impala[24] blue

jays[21], cotton top tamarin monkeys[25] red winged blackirds[26], pied flycatchers[27], and in the rat[28], these examples remain controversial.

To better understand the nature of the cognitive constraints and controversies that surround this concept of reciprocal cooperation it is important to understand the methods used to study it. The dominant approach taken by both evolutionary and psychological approaches to the study of reciprocity has been to employ game theory. Game theory is a collection of beguilingly simple 'games' that provide both a powerful means of quantifying the collective outcome of interacting individuals combined with a compelling theoretical framework upon which one can generate and test hypotheses about the cognitive processes of those individuals[5,7,20]. For a quarter of a century now, one particular game, the Prisoners Dilemma, has dominated both theoretical and empirical studies of reciprocity amongst unrelated individuals.

In the Prisoners Dilemma each player has two possible options, they may either choose to cooperate or choose not to. Choosing not to cooperate is termed defection. The resulting outcome each player receives depends upon both players' choices. Mutual cooperation results in a moderate reward to both players (Reward (R)) while mutual defection (Punishment (P)) yields a low reward to both players. The greatest reward is achieved through exploitation; by defecting upon a cooperator (Temptation (T)) and the lowest reward is received by the exploited player, one that cooperated with a defector (Sucker (S)). This relationship between the outcomes of collective choices, that T>R>P>S must be satisfied for a game to qualify as a Prisoners Dilemma[7].

A further concept, the Nash equilibrium, is required to fully understand why the Prisoners Dilemma was presented as such a powerful tool in the study of reciprocity . The Nash equilibrium is a complete analytical solution for social dilemma games that defines the action players *should* take when playing. The Nash equilibrium defines a collective act that neither player can improve upon by unilateral action. In the case of the Prisoners Dilemma, in a single round of play, regardless of the other player's choices, defection will always yield a greater reward. This is true regardless of whether one's opponent cooperates as one can exploit them by defecting, receiving the highest reward (T) or if they defected as one avoids exploitation and lowest outcome (S) in favor of a low reward (P). In essence, in a single round of the Prisoners Dilemma, each player *should* defect making mutual defection the Nash Equilibrium.

One might question the wisdom in studying cooperation in a game where the relationships between payoffs predict that each player should no do so. However, should players engage in multiple iterated rounds of the Prisoners Dilemma (iPD) a collective means of escaping this equilibrium becomes available. For a game to be described as an iPD a further rule must be conformed to; the value of two mutual cooperation choices (R) must be greater than the sum of the value of being exploited (T) and that of being exploited (S), or 2*R>T+S. This creates conditions where repetitive mutual cooperation exceeds that of unilateral or mutual defection. This elegantly formalizes the essence of what is being described by reciprocal cooperation where cooperation requires a short term cost, the risk of being exploited and forgoing the immediate benefits of

being an exploiter. However should such a cooperative act be reciprocated it provides each player with greater rewards over multiple iterations.

While the iPD may frame the tricky problem of reciprocal cooperation in an intuitive and quantifiable framework, it nonetheless makes cooperation a rather cognitively demanding task. For agents that wish to maximize their outcomes, the cost and benefits of cooperation must be compared. This requires some capacity of numerical discrimination and an ability to learn from prior outcomes. As cooperation in this context involves one forgoing a greater immediate reward for increased reward over time, it further requires that an animal does not devalue future rewards too strongly over immediate ones[7]. The rate at which future rewards are devalued is referred to as temporal discounting. As the future for any animal necessarily introduces uncertainty, it is proposed that such discounting, though at different rates, may be both universal among animals and under strong selection pressure[7,21]. This has led many psychologists to describe the iPD as primarily an extension of the temporal discounting problem[29], suggesting that cooperation in this context is as at least as much a cognitive feat as it is a reciprocal act.

The primary explanation for why reciprocal cooperation may be rare in animal populations is one of cognitive constraints amongst individuals of those populations. This reveals important assumptions that are made when applying a game theoretic framework to animal behaviour. In its pure form, game theory does not pose dilemmas to cognitively constrained individuals. Rather it asks how rational, emotionless geniuses should behave within particular games[5].

Rationality in this context makes the assumption that players behave in a normative manner by making individual choices in an attempt to maximize their outcome and do so by making optimal choices given the options available and are endowed with perfect knowledge of those options.[5]

Such assumptions of rationality present several problems when being applied to complex social phenomena. Classical game theoretical analyses predict that rationally self interested players will make decisions that align with Nash equilibria, however players, whether human or other animals, rarely play according to these strategies[5]. This may be in part due to the difficulty of ensuring that there are no aspects of an interaction that are being optimized by players that fall outside the scope of quantification ascribed by the chosen payoff matrix. Further, they are subject to the sorts of cognitive constraints detailed above, which limit their capacity to conform to economic predictions. In general animals have a range of reward driven behaviours that deviate strongly from economic expectations. One common example is that of Herrnstein's matching law. Matching refers to animals tendency to match their relative rates of response to the relative rates of reward[30]. For example if an animal is given two choices, once choice leads to a reward 70% of the time and the other 30% of the time. In such a situation the normative behaviour would be to choose the option where reward is available 70% of the time, 100% of the time, as this would yield the highest reward overall. However animals reliably match their ratio of choice to the ratio of reward, choosing the more rewarding option on 70% of choices and the less rewarding one 30% of the time. Again this points to another

limitation of economic rationality. In the context presented animals are behaving sub-optimally, however real world environments are characterized by uncertainty in both the sources and timing of rewards, suggesting that reacting to the statistics of foraging behaviour in such a manner may be adaptive and even more rational in general, even if it is not in the precise experimental context[31].

Game theory also defines cooperative choice as an *individual* act. Games such as the iPD demand not only that players behave as rational utility maximizers, but that they also do so in a purely self regarding manner. The payoff matrix that defines particular game formalizes the interdependence between collective choices and outcomes. Despite this interdependence off outcome the assumption remains that each player is trying to maximize their own outcomes and are entirely ambivalent as to the outcomes or behaviours of the other player(s). This assumption presents similar problems to that above, it is simply not an entirely accurate description of animal behaviour. Many species that are employed in cooperation experiment use 'safety-in-numbers' as a defense against predators, and such species often show a preference for proximity to conspecifics in safe cages or aquariums, even in the absence of a threat[3]. It has also been shown that rats[32,33], capuchins[34] and chimpanzees[35] are more likely to behave cooperatively when they can observe the behaviour of the conspecific they are interacting with. The nature of the social relationship between players can also have an important role upon the likelihood of cooperation. For example monogamous zebra finches have demonstrated an increased tendency to cooperate with their long term social partner compared to a novel

opposite sex conspecific, and this behaviour seemed to be, in part, determined by increased 'forgiveness' of occasional defection by their long term social partner[36]1. Status can also influence the likelihood of cooperative behaviour. In macaques, a despotic species, dominant individuals tend to be more cooperative while low status animals show the opposite behaviour[37]

These examples present powerful criticisms of game theory when applied to social behaviour. However there are further criticisms that can be leveled against the iPD in particular. One has already been mentioned; that arguably it is primarily an extension of the temporal discounting problem rather than a ecologically valid measure of reciprocity. The iPD also presents a stricter formalism than other social dilemma games as it requires a particularly specific relationship between the outcomes (2R>T+S) and that each player has the ability to exploit the other. This has led to rising skepticism as to whether it is an accurate representation of the type of real world cooperation that animals will actually engage in[7,21,33,38]. Or, more concisely researchers applying the paradigm are "searching in the dark for a cat that isn't there".[3]

Given that there are clear violations of the assumptions underpinning game theory and its application to animal behaviour, and of the iPD in particular, one might wonder what exactly is the appeal of such an approach, what exactly is it useful for?

---

1 This seems like an eminently adaptive trait in a monogamous species. It might even be a prerequisite.

2 Arguably this belief in humans as rationally unconstrained still defines libertarian economics and thought. The Randian focus on radical autonomy,

One important point is that these criticisms fall most strongly upon the assumptions of game theory in examining cooperation in the proximate sense. In the context of evolutionary game theory, these assumptions may hold more water. There is a strong analogy between the concept of rationality and the concept of inclusive fitness in evolutionary biology. While rational choice theorists assume that agents are making choices to maximize their utility, evolutionary biologist assume, and indeed in some cases can demonstrate, that animals will make choices that maximize their inclusive fitness. Evolutionary game theory, as it focuses upon the evolutionary processes, in the cases where animal behaviour is involved, is examining the success or failure of behavioural strategies or phenotypes within a population. The rational, emotionless decision maker described above in this case is natural selection rather than the animals themselves[39].

One can then proceed to turn this around, and look back again to the proximate context. Behaviours and the cognitive capacities that support them have undergone selection and are presumably adaptive in some manner. The implication of this is that natural selection may lead to behaviours that may approach rationality in one context but are not indicative of more general facultative competence. One possible illustration of this is that in autumn blue jays switch from consuming every acorn that they find, to caching them for later consumption throughout the winter, indicating an impressive control of temporal discounting in this *particular ecological context*.[7] In the context of an experimental iPD game, however, they only cooperate with a reciprocating partner when the burden of temporal discounting

is removed[21]. This suggests that while the behaviours associated with preparing for winter are under selection in such a manner that supports a greater control of temporal discounting, those relating to reciprocity are not. Another example may be that of the monogamous zebra finches mentioned above, as they appear to demonstrate cooperative behaviour in an iPD task only in a context where they interact with a mate. This, one could argue, may be adaptive in a monogamous pair bonding species as male and female partners work together during the breeding season to raise offspring[40]. When employing the iPD in the proximate sense, and examining function, one is essentially checking to see if selection has acted particularly upon the organism in question's tendency to cooperate.

In light of this it is interesting to note that humans while falling short of rational assumptions, in the context of cooperation, tend to fall up rather than down. Humans often tend to be more cooperative than an economic perspective might suggest, though many experiments have shown there is often a smaller fraction of subjects that behave in a self-regarding manner that is much closer to rational expectations[4]. If there is one clear 'arrow of history' in the study of human decision making, both economic and social, it points away from more rational and strategic assumptions towards more automatic, implicit and embodied ones. In an economic context the prevailing assumption for much of the last century was that humans obeyed the rational expectations of expected utility theory. The clearest attack on this belief came only in the late 70's when Kahneman and Tversky published their seminal paper on Prospect Theory where they clearly demonstrated the limitations of economic rationality in humans[41].

Similarly approaches to social cognition, particularly theory of mind, focused on rational hypothesis testing models and Machiavellian strategic analysis. Whereas recent work has focused on more automatic processes such as direct simulation and embodied inter-subjectivity[42].

This, now sharply declining, view of humans as rational and strategic[5]$_2$ has generated a view of social interactions as unboundedly complex in their predictive requirements. There is an assumption that a social interaction is much like a game of chess, each player trying to infer all the possible responses their opponent might make to their own strategic choices$_3$. From such a vantage point it is the rules of the game rather than the inferential ability of players that bounds the interaction. Similarly, game theory presents a context where the interacting agents are not constrained in their cognitive capacities, but rather by the nature of game they play. The Nash equilibria in a particular game is defined not by constraints upon the capacities of the player, but by their accurate perception and unconstrained consideration of all the strategic possibilities. One might suggest that the iPD in particular and game theory in general has defined cooperation, in part, because we thought that was the way we think in social and economic contexts. And, while humans tend to be cooperative in this game, we tend to do so in a manner that does not align with these preconceptions. Rather than behaving strategically

---

2 Arguably this belief in humans as rationally unconstrained still defines libertarian economics and thought. The Randian focus on radical autonomy, and the idea of 'letting the market decide' lean still strongly leverage this perspective.

3 It is telling that we developed computational approaches to defeat the best human players many years before we developed similar approaches to lift a chess piece in a similar manner to a human

and self-regarding, humans tend to be cooperative and other regarding[4,5]. More importantly, by casting cooperation in a strategic and individualized context it has de-emphasized ongoing social information, the implicit signals, other regarding behaviours and intrinsic rewards that are being seen as increasingly defining aspects of our real world social interactions[42,43].

The aspects of social interaction that game theory does not speak to how individuals make use of available social information to guide their social behaviour and decision making is described by the field of social competence. Social competence has primarily fallen under the auspices of the social sciences and has largely focused on the development and expression of social behaviours in humans[1]. Taborsky and Oliveira, however, have recently proposed a compelling approach to examine social competence in non-human animals. While the game theoretical approach we have explored has cast social behaviour very much in the context of cognitive evolution and ecology, Taborsky and Oliveira propose that flexible, adaptive social behaviour be observed through the lens of phenotypic plasticity. Phenotypic plasticity refers to the production of multiple phenotypes from a single genotype, depending on feedback from the biotic and abiotic aspects of their environments[44].

Social competence, from this perspective, refers to the ability of an animal to leverage evolved cognitive mechanisms that allow it to perceive the internal state of conspecifics and then to integrate this perceptual information to select the most appropriate social behaviour

from their repertoire of responses[4]. For example, rats appear to flexibly optimize their behaviour in response to experience in conflict situations by reducing fighting costs following a win or a loss. Previously victorious rats would defect against naïve conspecifics more quickly and with a lowered level of aggressive behaviour. Rats that had previously been defeated would concede more quickly and produced behaviour that reduced aggressiveness in their naïve opponent[1,45]. This example captures the manner in which animals can use experience to flexibly select an appropriate behaviour given their social experience and that they express this behaviour in a manner that naïve animals are sensitive to. As a result the potential cost of an agonistic interaction is reduced *for all participants*. Such use of social information has been observed in a wide variety of natural populations[1]

This discussion of cooperation began by referring to it, in the most general manner, as individuals "*choosing to interact in a manner where things work out collectively rather well*". The above example seems to fit this basic definition quite neatly. Both animals with experience of victory and those that experienced defeat were able to express behaviours as a result of this experience, that led to a decreased duration of aggressive or submissive responses required to resolve a conflict, that, if protracted could cost both interactors more. This type of agonistic interaction has rarely been cast in the proximate sense as an example of cooperation. Here winners reduced mutual costs by acting like a winner, in being more rapidly

[4] Though to what extent this is a result of predicting the behaviour of a conspecific and how much this requires active inference of that conspecifics internal states is left ambiguous by the authors.

aggressive, which reduced the duration of conflict required to lead to a resolution and preventing further escalation that could lead to injury[45]. It is quite possible that animals may engage in such behaviours in a manner that may also escalate conflict. Should a prior winner encounter an animal and misinterpret it's state a rapid onset of aggression could lead to an escalation of conflict. Similarly an animal that has previously lost may produce submissive responses that, if ignored by their opponent, may put them at a disadvantage. In principle each animal is incurring a cost, risking greater injury by their behavioural strategy, however if both interacting animals behave in a socially competent manner they both benefit from a clear display of their intentions.

I would argue that it is very human to balk at the thought of rapid aggression and submission in agonistic interactions as examples of cooperation. Surely cooperation is about working together, and so not being aggressive at all seems much more like cooperation than resolving conflict quickly. Yet such a position makes assumptions about the social structure within which a given organism exists and what behavioural options are within the perimeter of its behavioural repertoire. As such agonistic interaction seems to be a common component in various animal societies[1,46] and protracted conflict and the social stress that it results in often carries both proximate costs and impacts upon fitness in many species[47], then this particular display of social competence may be an instance of cooperation.

It is surprising then that Taborsky and Oliveira barely mention cooperation in their description of social competence. Rather they focus on animal's capacity to perceive relevant internal states of a

conspecific, and act upon this perceptual information to produce desired social outcomes. This presents more of a 'bottom up' approach to animal social behaviour, examining, within the ecological context and within the behavioural and perceptual repertoire of the organisms in question, their capacity for social competence. Game theory[5] on the other hand seems to adopt a top down approach, defining cooperation in more human, more economic terms and then searching for exemplars within animal societies. It is interesting that, asides from individual recognition, the primary cognitive requirements for reciprocity are numerical discrimination, temporal discounting, learning and memory. They are all essentially non social in nature[7]. Taborsky and Olivera, on the other hand, emphasize a very different view that social competence leverages a distinct sub-domain of general intelligence that is modulated by social experience[1].

And so we are presented with two approaches to social behaviour, one that provides a powerful mathematical framework and compelling formalism. However it suffers from simplifying assumptions and has little to say about the proximate manner in which animals interact, how they use the flow of ongoing social information to guide their choices in social contexts. It also presupposed a collection of sophisticated cognitive capacities[6]. While social competence focuses quite precisely on these aspects game theory ignores; adaptive behavioural plasticity in ecologically valid

[5] It is perhaps the more mathematically inclined, and/or those who rise to positions of influence within their field that tend to take this approach to social interaction.

[6] It may presuppose complete social competence.

social contexts. However this approach has little if anything explicit to say about cooperation, and lacks a mathematical framework. Integration of these two approaches, it would seem, may provide a compelling framework within which to examine the how animals engage in cooperation.

To proceed with such integration begs the question of how social competence can add to the concept of cooperation, and what might constitute a socially competent cooperative act. While the above example of agonistic interactions is a tentative example of cooperation, not all socially competent acts are likely to be cooperative. For example dominant Great Tit's (*Parus Major)* protect food resources they have monopolized by making alarm calls to deceive other dominant birds. But they do no do so towards subordinate ones. Subdominant animals behave similarly to both sub-dominant and dominant birds. These calls were modulated by food scarcity, or during periods where birds were more likely to feed[48]. This type of deception seems to carry the hallmarks of social competence. Animals here are monitoring the status of other conspecifics and mitigating the risks of both conflict, and the loss off monopolized resources. However as the deceiving animal is receiving most of the benefit at the expense of its potential competitor it would be hard to argue for this as a form of cooperation. This falls short of the example of agonistic interactions above[7] where each animal seemed to both accurately signal information relevant to the

[7] It is interesting, and perhaps mistaken that Taborsky and Oliveira refer to status, age and other differences as adding to the complexity of social interaction. For example when faced with a larger, higher status male, the decision of whether to engage in conflict seems less complex than when face with one of similar size and status. One probably should not.

interaction and respond to it appropriately and in so doing both animals appear to benefit.

The distinction between these two examples starts to circumscribe components of social competence that one might consider cooperative. At the core of this is the accurate signaling of intentions in a social context, rather than deceptive signaling, as this provides the other animal a greater opportunity to react appropriately, and may also risk exploitation. For such an act to be socially competent it would also have to based, to some extent, upon predictions of the other animals internal state and likely response[1]. I would therefore propose the following conditions for a socially competent cooperative act. First, each animal must honestly signal their intentions, in that they behave in a manner that makes them more predictable to a conspecific. Secondly each animal must attempt to interpret the signals of the conspecific they interact with and respond on the basis of the predictions they thus derive. Thirdly this response must increase the likelihood of benefits, or limit the cost of the interaction for both themselves and the interacting conspecific. Should these three conditions be met, I would argue that this constitutes a starting point from which to explore socially competent cooperative acts.

Fortuitously game theory offers one particular game, the Stag Hunt that formalizes quite precisely these three conditions. The Stag Hunt is in many respects very similar to the Prisoners Dilemma. Crucially it maintains a cost for cooperation, as to do so still carries the risk of receiving the lowest of all outcomes (S) if ones opponent defects. The Stag Hunt in does have one subtle but pivotal difference to the iPD. In the Stag Hunt mutual cooperation (R) offers the highest reward,

rather than in the iPD where exploitation (T), provides the greatest reward. This one simple change in the relationship between payoffs, whilst maintaining a cost for cooperation, drastically changes the dilemma the game poses. It means that the decision to cooperate is now contingent only upon the fear of being exploited whereas the presented double bind that combined a fear of exploitation with a temptation to exploit cooperators. Such exploitation yields an immediate high reward greater than that of mutual cooperation, and thusly required control of temporal discounting whereas the Stag Hunt relaxes this requirement. This is perhaps best illustrated by the fact that it has two Nash equilibria; mutual cooperation and mutual defection. Unlike the iPD where each animal is being asked to overcome defection based on the belief that their opponent will cooperate both now and in *the future* and whether they are able to overcome temporal discounting. The Stag Hunt, on the other hand focuses the dilemma on the current trial, and makes the best choice for each player the choice they believe the other is making. It first asks each if player can, coordinate their choices and then asks whether animals are able to collectively choose to mutually cooperate for greater reward. By posing this particular dilemma to animals with access to social information one has a potential means of quantifying cooperative social competence.

Having selected a game that emphasizes the role of social information, and that has previously been employed to do so, one is still left with assumptions of rationality that have been questioned in this introduction. It is important, and perhaps overdue, to make an important distinction. Game theory is the combination of two separate components, *games* and, of course *theory*. Theory provides

powerful simplifying assumptions, which, while useful, tends to reduce cooperation to an unintended by-product of the instrumental pursuit of private gain[49]. This poses a problem, as these assumptions are not likely to be met, and our interest is focused very much upon the social interdependence of actors. Games, on the other hand, simply propose strategic contexts. In so doing they precisely define the structure of interdependence between these actors. Here we intend to use theory to provide normative benchmarks against which trends in behavior can be compared while use the Stag Hunt game as a framework upon which social behaviour can be examined.

In the following chapters I will explore how rats integrate ongoing social information and rewards in a Stag Hunt task.

Chapter 1 explores in more detail the differences between the Stag Hunt and the Prisoners Dilemma and the assay that has been built to carry out the subsequent experiments.

Chapter 2 focuses upon social learning and examine the extent to which animals demonstrate an understanding of the economic terms of the Stag Hunt through observation of them interacting with an experimentally controlled stooge animal.

In chapter 3 we will focus on a series of experiments that examine the behaviour of freely behaving dyads in the Stag Hunt game and the role of social information

In the discussion we will return the idea of social competence mentioned here and examine whether the integration of these two distinct approaches to social behaviour presents a useful approach to

studying cooperation.

Chapter 1. The Stag Hunt

"*If it was a matter of hunting deer, everyone well realized that he must be faithful to his post; but if a hare happened to pass within reach of one of them, we would not doubt that he would have gone off in pursuit of it without scruple*"

Rousseau in a *Discourse on Inequality*

1.1 Social Dilemmas and Nash Equilibrium

The purpose in the following chapters is to examine the manner in which freely behaving rat dyads make use of social information to increase reward via collective action in an economic decision making task. It is the purview of this chapter to describe the genesis of the assay that has been designed for this purpose and to detail the theoretical considerations behind it. Both, the assay and theoretical approach that will be described here builds strongly upon prior work conducted within the lab and later published by Viana et al (2010) [28].

Viana *et al* (2010)[28] examined the behaviour of rats as they engaged in a particular social dilemma task, the Prisoners Dilemma. Game theoretical social dilemma tasks such as the Prisoners Dilemma are particularly useful for examining socio-economic decision-making as they provide a means of mathematically formalizing the relationships between collective action and outcomes. In so doing, they mathematically ground the problem of collective action and cooperation as a beguilingly simple mixed motive, two-person game with two choices: cooperation or defection. These choices intersect at four possible collective choices, mutual cooperation, mutual defection, unilateral cooperation, and unilateral defection abbreviated respectively as CC, DD, CD and DC. Each of these choices has an associated payoff: Reward (R), Punishment (P), Sucker (S) and Temptation (T) (see table 1.1). A social dilemma game here is defined as one where mutual cooperation provides the best collective outcome, but where this outcome is constrained by the temptation to cheat (T>R) or the risk of being cheated upon (P>S). [20]

	Cooperate	***Defect***
Cooperate	***R,R***	***(S,T)***
Defect	***(T,S)***	***(P,P)***

As the title of this chapter suggests; a different social dilemma game, the Stag Hunt was employed in subsequent experiments rather than the Prisoners Dilemma. There are subtle but important differences between these two games that motivated this choice and that we will now explore. In the Prisoners Dilemma the outcomes are ordered thusly T>R>P>S, placing both, the fear of cheaters (P>S) and the temptation to cheat (T>R) as constraints upon cooperation. The Stag Hunt outcomes are ordered thusly (R>T=>P>S). By having the temptation outcome lower than reward (R>T) the temptation to cheat (defecting upon a cooperator) is removed, but the risk of being cheated upon is maintained by keeping punishment greater than sucker (P>S). This simple distinction between the Prisoners Dilemma and Stag Hunt outcomes leads to fundamental differences in the dilemma posed to players of each game. These differences are perhaps best described by comparison of the Nash Equilibria[5] of each game.

Nash Equilibria are complete analytical solutions derived from the relationships between the outcomes defined by each social dilemma game. A Nash equilibrium, simply put, is a collective choice from which an individual could not do better by changing its strategy[5]. In the Prisoners Dilemma the double bind of the fear of being cheated (P>S) and the temptation to cheat (T>R) means that mutual defection is the single Nash Equilibrium and defection the dominant strategy[20]. The Stag Hunt on the other hand has two Nash Equilibria, mutual cooperation(R) and mutual

defection(P). The mutual cooperation equilibrium is considered payoff dominant, in that it offers the highest outcome whilst mutual defection is considered risk dominant, as it provides a lower positive outcome but mitigates the risk of receiving the lowest sucker outcome[50].

To create an intuition for this, consider two players with complete knowledge of the outcomes engaging first in a single instance of the Prisoners Dilemma. Now imagine you are player one and player two has announced that they will definitely cooperate. As defecting on a cooperator yields a better outcome than cooperating with a cooperator (T>R) you should defect to get the maximum outcome. Now let us assume that your opponent will certainly defect, the best choice here is that you also defect as cooperating with a defector yields the lowest outcome (P>S). As defection in the Prisoner dilemma yields the greatest outcome regardless of the choice of one's opponent, it is considered the dominant strategy and mutual defection forms the single Nash equilibrium. Now if we follow the same reasoning with Stag Hunt, we are left with a very different choice. If one's opponent announces they will cooperate, cooperation becomes more lucrative as (R>T) and if they will defect then defection becomes the better choice as (P>S). Consequently in the Stag Hunt there is no dominant strategy and there are two Nash Equilibrium; mutual cooperation and mutual defection[50].

This difference in Nash equilibria in the Prisoners Dilemma and Stag Hunt creates very different predictions for what players *should* do in a single round of the game. In the Prisoners Dilemma players are expected to defect. In the Stag Hunt a player should make the same choice as they expect their opponent will make. However, expanding these games from a

single choice it is possible that in each case, there are conditions where two players may engage in consistent cooperation

The concept of the Nash Equilibria as described above has limitations in its formal applicability in the iterated case, where agents engage in *n* sequential rounds of choice. In the iterated case, the Nash Equilibrium can be calculated over multiple trials, and thusly incorporate the history of outcomes; this can lead to novel Nash equilibria over different strategies. Nonetheless, the Nash equilibrium concept, in the manner in which we have described above, captures a clear description of the fundamental dilemma each game poses to its players. A further rule is applied to create the iterated case of the Prisoners Dilemma (iPD). Here the repeated mutual cooperation is fixed as a greater outcome over multiple trials by fixing the reward outcome as greater than the sum of a temptation and sucker outcome so that 2*R>T+S. This ensures that mutual cooperation is more lucrative over multiple iterations than alternating defection and cooperation. This provides agents a means of escaping the Nash equilibrium through collective cooperation assuming that they have memory of at least one iteration back and have sufficient control over temporal discounting. Temporal discounting refers to rate at which a delay in receiving a reward devalues it in comparison to those received directly. Should both the criteria be satisfied, it is in principle possible that subsequent mutual cooperation will be considered beneficial over proximate temptation.[7] In the iterated Stag Hunt (iSH), the Nash equilibria remain the same so that in each iteration, each player's best choice remains coordination where each player makes the choice that they believe their opponent has made. Across multiple iterations, the question now asked is whether interacting agents can collaborate to achieve more coordinated

cooperation trials than coordinated defection and thusly increasing the proportion of R than P outcomes.

The concept of Nash equilibrium finds further limitations in its applicability to a social decision making task in the rat. Primary amongst these is that for such rules to be entirely applicable each player must have complete knowledge of the outcomes available to themselves *and the other player* prior to making a choice. The iterated case of each game provides a means for rats to learn the pattern of choices of the other. This provides each player with a means to develop expectations of their opponents' choices and how they might influence received outcomes. This may provide a functional approximation of the outcomes that the other receives. However, the Nash equilibrium concept offers analytical solutions to games. Its purpose is to demonstrate under various conditions what players *should do,* not predict what players will *actually do*[5]. More precisely, the Nash concept cannot make precise predictions about the selection of super-game equilibria, that is, about the outcome of ongoing, mixed motive games. More importantly, it has little to say about the dynamics by which players can move from one equilibrium from another[20]. The Nash concept also has very little application in how agents' behaviour in a game theoretical context may be based upon learning[5]. Nonetheless they do provide a clear description of the fundamental dilemma each game poses to its players. In the context of iterated versions of these games, for the Prisoners Dilemma, the primary question being asked is whether two players can overcome the dominant strategy of defection and maintain mutual cooperation. The iterated Stag Hunt (iSH) asks first if animals will coordinate, then asks whether they do so via mutual cooperation or mutual defection.

1.2 Choosing a game

While the Prisoners Dilemma has been more widely applied and is generally accepted as the 'gold standard' for measuring reciprocity, there are several theoretical and practical reasons why the Stag Hunt provides a compelling alternative. Firstly, the Prisoners Dilemma is being increasingly seen as a very particular case, the conditions of which are rarely met under natural conditions, questioning its empirical relevance[35,40,51]. The Stag hunt, in contrast may be a more accurate description of the sort of social dilemmas animals may face in real world situations[35]. The Stag Hunt describes situations where individuals must collaborate with others, that the benefits of this collaboration are greater than solitary action, but that cooperation risks loss of those benefits. This has been proposed as an accurate description of collective hunting in a number of cases in chimpanzees[35,38] and proposed as such for early human foraging[35] and orchestrated whale hunts[52].

This difference has important implications for attempting to understand the motivations behind a player's choice to defect in each game. The Prisoners Dilemma offers two motivations for defection, while the Stag Hunt suggests only one. In the Stag Hunt, defection is expected when a player doubts their opponent will cooperate. This also motivates defection in the Prisoners Dilemma, but in addition to this there is also the temptation to exploit a cooperator for greater reward. This mandates that one take a further inferential step in the iPD than in the iSH when attempting to determine the motivations behind the propensity to defect.

While the iSH does provide a less complex dilemma for its players, it does not provide a less interesting one. In the iSH, by ensuring that the sum of the two possible outcomes for cooperation are less than that of the two possible defect outcomes so that R+S<2x(T|P), one ensures that a self interested player will only cooperate if they expect the other to do so. This is because if one plays against an unpredictable and random player, on average, choosing to cooperate will yield less reward than the constant reward received for defection. The cost of cooperation becomes a probabilistic one, choosing to cooperate carries with it the risk of receiving the lowest payoff. This captures a fundamental component of reciprocity; that cooperation can be costly. However, unlike the iPD where cooperation only increases rewards only over multiple iterations (T<P), in the iSH, the benefits and cost of cooperation (R or S) are received in the same trial. While relaxing the requirement to minimize temporal discounting, this fails to satisfy a further requirement of reciprocity; that there be delay between the incurring of a cost and the resulting benefit.

This is not always the case. The Stag Hunt has also been used to explore mutualisms in a variety of species and contexts[35]. This is not incorrect; the Stag Hunt can indeed be used to formalize a mutualistic interaction, where there is no cost for cooperation. This is achieved by simply increasing the R outcome so that it is greater than 2x(T|P). This means, in the iterated case, if one is playing against a random player, cooperation will yield a greater reward on average, and thus remove the cost of cooperation. Arranging the Stag Hunt payoffs so that cooperation is risky and therefore incurs a cost arguably means that it formalizes a form of cooperation that lies in between mutualism and reciprocal altruism.

Here we are attempting to understand the manner in which freely behaving animals collaborate to achieve reward. Approaching this in the context of a formalism that is more likely to describe real world interactions potentially adds to the empirical relevance of any findings. The Prisoners dilemma adds a further layer of complexity and cost to cooperation over the Stag Hunt but offers the opportunity to examine it in the classical context of reciprocal altruism. While Viana *et al* (2010) have demonstrated that individual rats are capable of cooperation in the iPD task when faced with a reciprocal opponent, freely behaving animals have not demonstrated this capacity. While this might well be possible, our question here is not simply *will* animals cooperate under these conditions but *how* they integrate outcomes with an ongoing flow of social information. It therefore seems pertinent to approach the question of cooperation under conditions where such behaviour is more likely to occur, its explanation more easily approached, and where the choice to cooperate still poses a dilemma. In this case, that dilemma is defined by risk and the ability to determine what one's opponent will do in the current trial rather than in the future.

1.3 Implementing an Iterated Stag Hunt Task.

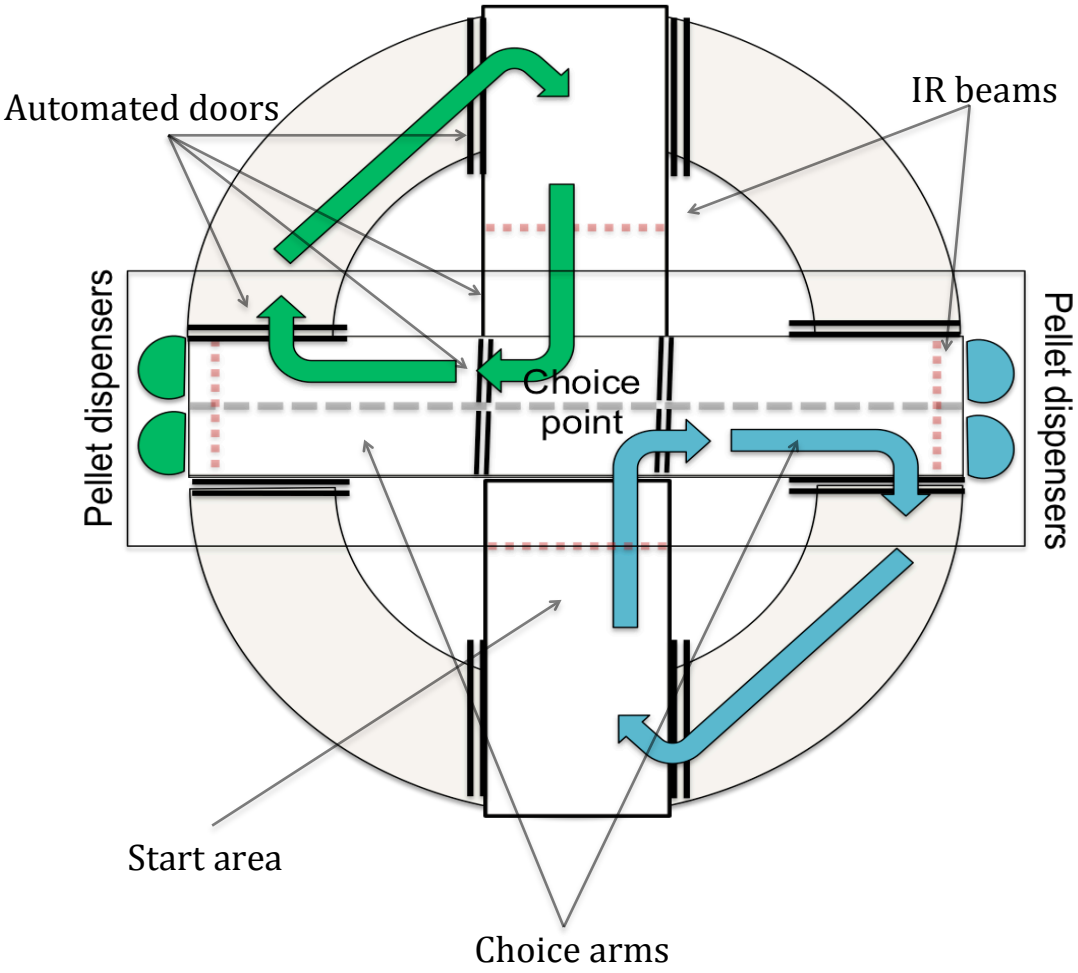
Viana *et al* (2010) implemented their version of the iterated Prisoners Dilemma (iPD) in the form of a dual two alternative forced choice task (2AFC). In this example one compartment of a T-maze was defined as cooperate, the other as defect for a freely choosing animal and the same for an adjoining T-maze where the experimenter controlled the choice of a stooge animal. This instantiation of the iPD required an experimenter to manually secure the rat's choice during each trial once it had shown a preference for either the cooperate or the defect compartment,

deliver outcomes, and to return the rat and stooge to a start location following each trial. Similar versions have been implemented with levers[53]. To reduce the role of the experimenter and in an attempt to increase the number of trials performed by rat dyads, we designed and implemented a fully automated dual 2AFC task along similar principles.

This particular design of the assay has a further advantage. Rats live in burrow systems consisting of various chambers interconnected by tunnels. In such environments, subordinate males have been shown to coordinate defensive behaviours, collaboratively blocking access to tunnels that lead to food resources against larger, dominant males[54,55]. This design of the assay echoes aspects of this ecological context, consisting of larger chambers and interconnected corridors, and defining cooperate and defect as the end point of a trajectory towards one spatial location or another.

This design consists of two identical, independent T-mazes (70 cm x 40 cm, 30 cm high each individual maze) that were placed together for the iSH task (see schematic 1.1). Separating these two mazes is a single transparent perforated acrylic wall that allows each animal to see, smell and make limited contact with the other. Automated infra-red beams (dashed red lines) were strategically placed in each maze to record access to either choice arm of the T-maze and when animals return to the start area. Automated doors controlled access from the start area to a central decision area, to the choice arms and from the return corridors back to the start area. This series of doors ensured that the progression of one or both rats from start area, to the choice arms at the end of which rewards were delivered and then back to start area could be strictly controlled. The

direction of movement is indicated by the colored arrows in (schematic.1.1) Automated pellet feeders placed at the end of each choice arm delivered pellets. The activity of these doors were controlled using Graphic State 3.03 software and the Habitest interface (Coulbourn Instruments Allentown, PA, USA).



Schematic 1.1 | Maze design
Design of maze. Red dashed lines indicate infra red beams. Dashed grey line indicates perforated acrylic partition, double black lines indicate automated doors. Green arrows indicate the direction of movement for cooperative choice and blue indicates the direction of movement for a defection choice.

1.4 Determining a payoff matrix

The relationship between outcomes in the iSH game is organized as R>T>=P>S. Our interest here is to determine a set of outcomes that are both discriminable to the rats and the rewards as low as possible to increase the number of trials rats will complete within a session. To determine this, we conducted a series of simple discrimination experiments where preference for different outcomes were tested pairwise with each of two rewards being available consistently on one of two arms of a single T-maze. All rats were tested individually.

1.5 Animal Subjects

The experiments were performed using 12 pairs of male non-litter mates of the outbred Sprague Dawley rat strain, from Charles River, Barcelona, Spain. All animals were housed in pairs under 12 h light/dark cycle. Experiments were conducted during the light period. All rats were habituated to the experimenter alone for one week and for approximately 4 days to the automated double T-maze assay and to the novel food used for the positive reinforcements in the iSH stooge task. All rats were given free access to food and water.

1.6 Discrimination Results

In each of the following early experiments, four naïve animals were each exposed to one of three different single pellet discriminations: 1-0, 2-1 and 3-2. In each of these cases, we examined the proportion of total choices made to the higher arm side. Rats in all cases demonstrated a preference greater than chance for the more rewarded arm (one tailed t

test: 1-0 t_{19}=6.557, 2-1 t_{19}=7.208, 3-2 t_{19}=7.5526, $p<0.001$). These results indicate that different outcomes ranging from 0-3 pellets in single pellet increments are both discriminable and desirable to rats.

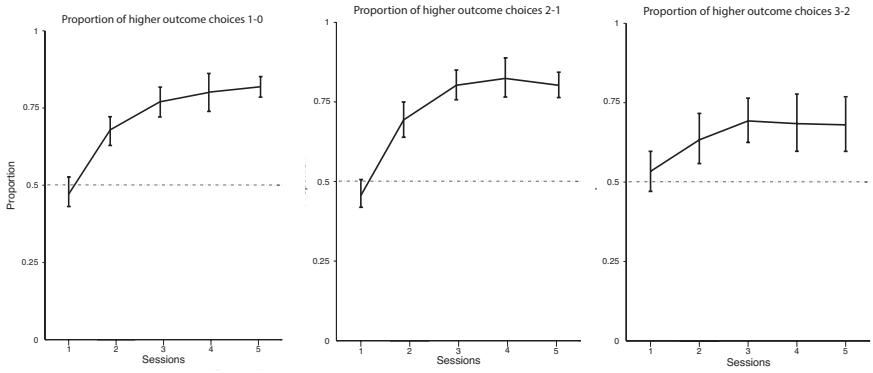


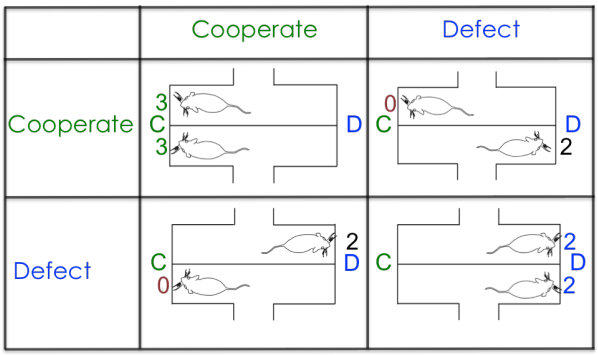
Figure 1.1 | Rats discriminate single pellet differences from 0-3
Proportion of total choices made to arm of maze with the higher reward for 1-0 pellets (a), 2-1 pellets (b) and 3-2 pellets (c). In each case significant preference for the higher reward arm is revealed by the second session.

1.7 Building the iSH Payoff Matrix

Having established that rats can both, discriminate between 0 and 3 pellets in single pellet increments and demonstrate a preference for the larger reward in each case, we can now use these pellet quantities for each of the outcomes in the iSH game. This allows us to create the payoff matrix where R=3, T=2 P=2 and S=0 as shown in table (1.2) below and instantiated in the context of dual T-maze 2AFC task as indicated in figure 1.2

(1.2)

	Cooperate	***Defect***
Cooperate	***3,3***	***(0,2)***
Defect	**(2,*0*)**	***(2,2)***



Aligning the payoff matrix in this manner allows us to emphasize and leverage the risk component of the iSH game. This organization of the two outcomes for cooperation so that R+S<2x(T|P), ensures that a self interested player will only cooperate if they expect the other to do so, as mentioned above. Organizing the payoff structure so that defection provides a constant moderate outcome that is only exceeded by collective cooperation allows us to make the following important statement. Freely behaving rats engaging in the iSH task that understand the payoff matrix, will only engage in cooperation at a level greater than chance if they are able to do so collectively.

1.8 Discussion

Thus far, we have explored the characteristics of the iSH and iPD as candidate games to explore the socio-economic decision making of rats. The iSH was chosen as a more appropriate game for freely behaving rats to engage in as it provides a more interpretable dilemma that may also be a more representative formalism of real world interactions that animals may actually engage in. We then proceeded to describe the design of the assay, a combined double T-maze, and via simple discrimination experiments determine reward values that satisfy the iSH payoff matrix.

The iSH was considered to be more interpretable than the iPD as in the iPD an animal may defect out of fear of receiving a low S outcome or because they prefer the highest T outcome. In the iSH, each choice should be motivated by what each player believes the other player has chosen. Because R provides the highest reward immediately, it removes the need for temporal discounting, but further, it emphasizes that for each player each choice is based on the perception that the other rat will choose the same. By choosing a payoff where $R<2*T|P$, we emphasize the risk aspect, adding a cost to cooperation which captures an important component of reciprocity, that cooperation be costly.

By performing simple discrimination experiments comparing 0,1,2 & 3 pellets, we were able to select a payoff matrix that satisfies the requirements of the iSH and presents outcomes that are both discriminable and desirable to rats. One possible concern here is that it limits the dynamic range of outcomes. Rats will only gain a single pellet per trial through mutual cooperation, but a loss of two for unilateral cooperation.

While this preserves the emphasis on risk discussed above, and, rather than having larger reward differences, increases the likelihood of animals engaging in more trials per session before becoming satiated, it does decrease the likelihood of sustained cooperation between freely behaving animals.

Another benefit of the assay above is that it is remarkably versatile in its applicability to social dilemma games. By simply changing the payoff matrix, we can effectively instantiate the iPD or other games we have not yet discussed such as Matching Pennies or the Snowdrift, or any distribution of outcomes that an experimenter should deem useful. Each maze is a separate module, which means that we can occlude visual information, or even separate the mazes entirely so that all social information is removed. As this design is fully automated, it also permits the control of one or both animals' behaviour by limiting access to one or other arm of the maze or forcing one or other animal to choose first.

Chapter 2: Stooge Stag & Hare

You go your way
I'll go your way too.
Leonard Cohen 'Book of longing'

2.1 Introduction

Rats are social animals. A growing collection of work continues to demonstrate that rats can make use of proximate social information to learn about their environments[56]. Such social learning strategies can be broadly split into two categories. The first and perhaps simpler mechanistic category is social facilitation, where the presence of a conspecific has an 'energizing effect' in that the presence of social contact arouses general drive and increases the likelihood of behaviours, such as exploration, of interacting animals in general. The second, and for our purposes more relevant category, is the various forms of observational learning where animals are able to make specific inferences about their environment from a social interaction[57].

These types of social learning strategies have historically been predominantly approached through various demonstrator-observer paradigms. This methodology employs a demonstrator animal that displays information to an observer animal whose behaviour is then examined for evidence of responses to the information that has been transmitted[58]. This demonstrator-observer paradigm has been used in a myriad of different contexts to demonstrate an expanding portfolio of sensitivities and responses to social information. A classic example of such directed learning through social interaction is the transmission of food preference, where observer rats learn a preference for novel foods that a demonstrator conspecific has recently ingested via olfactory cues present in the breath of that animal[59],[60]. Rats also show a similar potentiation of stimulus action reward associations through observation of operant behaviours. This has been demonstrated in multiple cases where rats learn more quickly to

interact with some manipulandi, such as a lever or joystick, to achieve reward after having observed a conspecific do so[58,57,61].

These directed sensitivities and responses to social information in the rat have also been demonstrated to be modulated by prior experience. Interestingly this has been predominantly shown for aversive rather than appetitive stimuli. For example, rats copy the threat avoidance responses, freezing, when exposed to a freezing demonstrator. However, only animals that previously had prior experience to a similar noxious stimulus display this behaviour, whereas those that have not had this experience do not[62,63].

Rats also appear to have a notable capacity to act in accordance with expressed preferences of demonstrator animals in the absence of obvious self-benefit. Prior work in the lab (see appendices) has shown that rats will respond to the expressed preferences of demonstrator for one arm of a modified version of the T-maze design detailed in the prior chapter by providing them access to a rewarded rather than unrewarded choice[56]. This conveniently indicates that in an experimental context very much like the one we are about to explore, observer rats are sensitive to expressed preferences of demonstrators for one option over another and can respond accordingly.

Each of these cases involves a sort of copying, that the observer animal increases the likelihood of certain responses that match those behaviours of a demonstrator animal. This can be responding to food seeking behaviour by choosing an arm that a demonstrator signals preference for, or being more likely to engage in threat avoidance responses when they are also displayed by a demonstrator. Defining the

type of cognitive mechanisms that are supporting these behaviours is notoriously tricky[14,61]. Each of the cases above can be, at least in part, explained by what are referred to as local and stimulus enhancement. The precise distinction between these two categories of social learning and indeed the extent of their borders remains only partially defined[14,61]. Nonetheless together they refer to the increased likelihood of the acquisition of behaviour towards or in response to some stimulus, object or a location through the observation of a demonstrator performing an action upon that object or within that location[61] . These deceptively simple social learning strategies can be surprisingly powerful and are often sufficient to explain foraging behaviours that might at first glance seem to depend upon more complex mechanisms[64].

These forms of social learning, and quite possibly the majority of others forms, depend upon two, in principle, separable learning processes, on one hand, social and on the other hand, instrumental. In the case where animals learned more quickly to press a lever to obtain a reward via observation of a demonstrator doing the same, social information seems to orient the observer towards the lever. Upon interaction with the lever, instrumental processes then drive learning[61]. In the case of social transmission of fear, it appears that individual prior experience is combined with social information to produce freezing during interactions with a demonstrator animal that displays threat avoidance responses.

While rats are proving a valuable model for studying social behaviour, they have been, of course, more extensively studied in non-social contexts. The type of socio-economic interaction that concerns us here, formalized by game theory, emphasizes the interaction of both social

and non social cognitive processes, particularly those that are defined under the umbrellas of value based or perceptual decision making. Decision making here refers to the ability of animals to choose between competing courses of action based on the relative value of their consequences and available perceptual information. Rats have demonstrated important capacities to engage in such decision making, such as transitivity[65], speed accuracy trade offs[66], probability estimation[67], control of temporal discounting[28] and the ability to act upon their degree of belief or confidence in a decision[68].

Recently, work by Tervo *et al* (2014) examined the non-social aspect of rats' decision making capacities in a game theoretical context[69]. Here, agent animals played a game called matching pennies. In the matching pennies game, one player is rewarded for making the same choice as the other, while the other player is rewarded for making a different choice than its opponent. This creates a context where each player should attempt to avoid exploitation by behaving unpredictably[69]. In this context, agent behaviour was examined without the presence of conspecifics or other social cues. Rats were instead pitted against an algorithm that increasingly approached true randomness, removing any obviously social component from the task. Rats in this task showed a capacity to counter-predict their algorithmic opponent and engage in strategic model based behaviour. Further, when it was not possible to counter predict the algorithm strategy, animals then adopted a safer 'stochastic' mode where outcomes were ignored.

Decision making is a fundamentally integrative process which hinges upon the capacity to knit together the causal relationships between

actions and consequences with the stimuli that predict them[70]. Approaching decision-making in a social context adds social signals transmitted by a conspecific to this already complex melding of action and outcome. The demonstrator-observer paradigm is also a ubiquitous way of approaching game theoretical social decision making in the rat, as it is in a variety of other species. In this context, however, the language changes, the demonstrator becomes the stooge animal and the observer, the agent animal.

Viana *et al* (2010) as was extensively referred to in the prior chapter, employed precisely this approach. By controlling the behaviour of a stooge animal, they were able to quantify the strategic responses of the agent rat to strategies of the stooge animals it faced[28]. This indicated that the added complexity of adding a conspecific to the complex cognitive mix of decision-making was not prohibitive of model based or strategic decision-making.

An important point in Viana *et al's* study is that the approach taken, despite the use of conspecific, may still emphasize elements of non-social decision making rather than social learning strategies. On each trial, the position of the stooge animal dictated the outcomes from which the agent animal could choose. Crucially the stooge's behaviour was predetermined to be consistent with particular strategies, either random or reciprocal. The agent animal was being observed for its capacity to respond cooperatively in the face of such strategies. The stooge animal, asides from its location, was not required to signal any information that was explicitly social, such as a defensive posture or a preference for one or other choice that the agent could make. The stooge animal may be

effectively no different from a cue. It is quite possible that more social elements such as local or stimulus enhancement, or social facilitation played a specific role in the behaviour of observer animals, but there is no evidence to that effect.

A fundamental aspect of the project this thesis describes is to enact a radical departure from the observer-demonstrator paradigm and instead examine the behaviour of freely behaving animals as they engage in social decision-making. The type of strategic decision described by Viana *et al* and Tervo *et al* is of central interest. The first step in this process, described in this chapter, is to determine the extent to which animals can integrate reward history and social information under the payoff conditions of our game of interest; the Stag Hunt.

To quantify the extent to which rats can effectively associate the contingencies between their choice, their opponents, and the outcome, we will initially make use of the stooge-agent paradigm by controlling the behaviour of a stooge rat and observing that of an agent rat. As was detailed in the prior chapter, the Stag Hunt differs from the Prisoners Dilemma in that it does not have a dominant strategy and the optimal choice is to do what one believes the other has done. Coordination, making the same choice as one's opponent, is therefore the choice that yields the greatest outcome on each trial. One means of measuring a rat's capacity to associate these contingencies is to have a stooge animal compelled to choose to either cooperate or defect and to do so first so that its location is clear to the agent animal before it makes its choice. Then one can measure the extent to which an agent animal coordinates with the stooge, and thus receive the greatest outcome on that trial.

The findings of Viana *et al* and Tervo *et al* inform this methodology in two important ways. Both, in the presence,and absence, of social information, respectively, rats in these experiments demonstrated a proclivity for responding to the strategy of their opponent. This suggests fixing the stooge's choice as pseudo-random so that on any given trial, the only information an agent rat can accurately make their choice upon is the stooge's current position and the resulting available outcomes.

Tervo *et al*'s work also suggests a further tentative prediction. They found that rats that were confronted with an algorithm that they could not counter-predict or effectively model, adopted a safe stochastic decision mode. Now, the Stag Hunt differs from the matching pennies game in that defection offers a safe constant reward. This implies that an agent animal, faced with a random stooge which prohibits strategic prediction and association of its behaviour effectively with outcomes, should show an increase in 'safe' defection choices.

2.1.2 Experimental Environment

The behavioural apparatus consisted of two identical fully automated individual T-mazes that are placed together for the iSH task. A schematic view was shown in the prior chapter. Custom made acrylic mazes (70 cm x 40 cm, 30 cm high each individual maze) (Gravoplot, Sintra, Portugal) were automatically controlled using Graphic State 3.03 software and the Habitest interface (Coulbourn Instruments, Allentown, PA, USA). Custom-made automatic doors (Champalimaud Scientific

Hardware Platform, Lisbon, Portugal; Mobiara R&D Design, Lisbon, Portugal; WGT-Elektronik, Kolsass, Autria) triggered by infrared beams (Champalimaud Scientific Hardware Platform, Lisbon, Portugal) controlled the position of the animals in the mazes. Each T-maze had a central corridor as starting point (Fig S1A), and two lateral choice arms at the end of which there was a food magazine that delivered food pellet rewards. For the duration of the experiments, one arm was fixed as cooperate and one as defect. Once both rats had entered either one of lateral cooperate or defect choice arms, each rat could retrieve the appropriate food rewards (45 mg pellet, reference F0021, BioServ, Frenchtown, NJ, USA). This then triggered the simultaneous opening of the appropriate 'reward' and 'start box entry' door, for each animal, after 3 seconds. This leads to the closure of the access doors until both rats return to their respective starting points, thus initiating another trial.

2.1.3 Animal Subjects

The experiments were performed using 27 pairs of male non-litter mates of the outbred Sprague Dawley rat strain, from Charles River, Barcelona, Spain. All animals were housed in pairs under a 12 h light/dark cycle. Experiments were conducted during the light period. Before starting the experiment, stooge and agent rats were habituated to the experimenter and to the food rewards used in the iSH task for one week and to the automated double T-maze assay for approximately 10 days. Each experiment used naive cage-mate pairs, one of which played the role of the stooge and other the agent. All rats were kept at 95% or higher body weight and while not food deprived, were given access to a limited number

of 14 standard chow pellets after each session. They were given free access to water.

2.1.4 Training

Here the training sequence for agent rats and stooge rats was the same. Firstly rats were handled for one week in the experimental room and introduced to the food rewards. Second they were gradually introduced to the iSH assay. The iSH assay is a complex environment with automated moving doors and pellet feeders that limit the movement of rats within the assay, in particular, the order and directionality of their choices. The assay, primarily due to these automated doors, is initially mildly anxiogenic to the rats, and thusly they are introduced to the maze with all doors open. Over the ten days of habituation to the environment, they are introduced to these doors, during this period they also learn the location of pellets delivered by the pellet feeders. Once each rat moves comfortably around the maze with all doors active and receives a single pellet reward for each completed trial, displays no significant bias to either side, they are considered ready to start the experiment.

2.1.5 Statistical Analysis

All the analysis was performed in Matlab 2010 (The 7.10.0 Mathworks, Inc) mean and standard error of the mean are reported. All data satisfied assumptions of normality unless otherwise mentioned.

2.2.1 Coordination Experiments.

Here we examine the behaviour of 15 stooge, agent rat pairs as they engage in a standard iSH task. On each trial, either the cooperate or defect arm was made accessible to the stooge animal, while the other remained closed. Once the stooge has accessed the available arm, doors leading to the cooperate or defect arms of the maze opened providing agent rats the ability to freely choose one or the other arm. Agent rats receive rewards consistent with payoff matrix shown in (table 2.1) while stooge rats always receive a single pellet.

Table 2.1

	Cooperate $_{agent}$	***Defect*** $_{agent}$
Cooperate $_{stooge}$	1,***3 (R)***	1,***2 (T)***
Defect $_{stooge}$	1,***0 (S)***	1,***2 (P)***

This allows us to determine the extent to which animals understand the contingency between their choice, the stooge's position and the outcome and not upon outcome statistics resulting from strategic behaviour on the part of the stooge. This iSH payoff task rewards coordination, choosing the same arm on each trial that the stooge had accessed. The coordination was a measure of performance. The cooperation arm provided the highest outcome (R) only when chosen in coordination with the stooge, while the defection choice always provides a moderate constant outcome regardless of the stooge's location. Consistent defection in this task would suggest that the agent animal has not learned the association between their choice, the stooge's choice and the outcome,

as they would be choosing a constant moderate reward even when the stooge has provided access to a higher reward on the cooperate arm.

2.2.2 Coordination Experiment Results

Agent rats performed approximately 50 trials per session (mean=53 trials ±0.45) and sessions lasted approximately 40 minutes. Fig 2.2.1 shows a representative raster of the raw trial by trial choices to cooperate (green) or defect (blue) of both stooge and agent and whether the choices were coordinated (white) or anti coordinated (black).

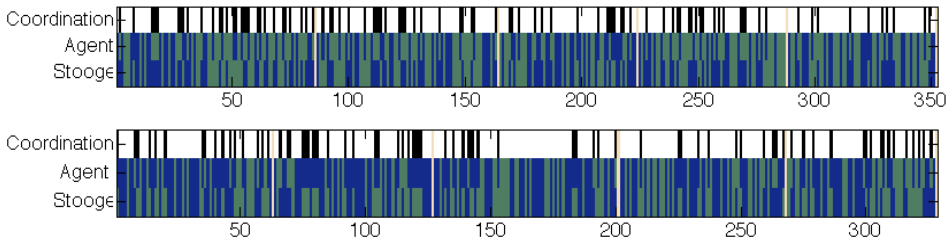


Figure 2.2.1 | Examples of rat and stooge behaviour.
First row shows coordination (white) and anti coordination (black) between cooperate (green) and defect (blue) choices made by agent (second row) and stooge rats (third row) for all five sessions, sessions boundaries in beige. Data shown for two representative stooge agent pairs for all trials and sessions.

The central question in this experiment is whether agent animals display behaviour consistent with an understanding of the contingency between their choice, the stooges and outcomes. The primary measure of this is the extent to which they coordinated with the stooge animal. Across all sessions and animals, we see that they coordinate robustly greater than chance (one sample t_{74}=13.4694, ±0.12 $p<0.0001$). No significant change was seen over sessions (fig 2.2.2a). Agent rats show a robust preference and capacity for coordination. Yet, despite having access to the stooge's choice before making their own, they do so only moderately (mean=0.67

±0.06). This may be due to animals developing a preference for either unilateral cooperation or defection.

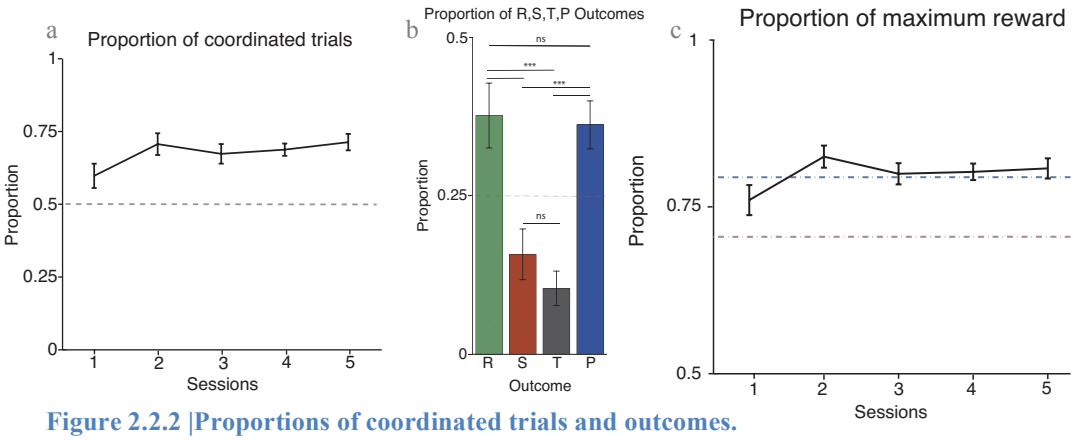


Figure 2.2.2 |Proportions of coordinated trials and outcomes.
Mean proportion of coordinated trials by session (a), outcomes (R:green,S:red,T:grey,P:blue) across all sessions (b), (c) shows the mean proportion of maximum reward (provided by complete coordination with the stooge) received over sessions.

To determine whether there was a systematic bias in agent rats' choices, we examined the distribution of outcomes. It is important to remember that the stooges' choices were fixed to a pseudo random distribution and therefore cooperated on approximately half of all trials, and defected on the rest. A systematic bias would therefore be indicated by a preference for either coordinated cooperation or defection or anti-coordinated cooperation or defection. Figure 2.2.2b shows the proportion of outcomes received for coordinated choices R and P and anti-coordinated choices S and T. A repeated measures ANOVA revealed a significant effect of outcome upon proportion of choice ($F_{3,72.12}$=0.7312 *p<0.0001*), post hoc analysis by Tukey Cramer HSD indicated that the mean proportion of R outcomes (M=0.3396 ±0.0110) and P outcomes (M=0.3314 ±0.017) were equivalent and both were greater than anti-coordinated outcomes (T:M=0.1574 ±0.0105, S:M=0.1714 ±0.0101),

between which there was no significant difference. Both anti-coordination outcomes were significantly lower than chance (one sample t, S: t_{74}=-7.5409 $p<0.001$, T (t_{74}=-9.104, $p<0.001$).

This suggests that the proportion of anti coordination trials is not motivated by a consistent bias for cooperation or defection. Another measure of performance, and one that that is likely more important to the agent rats, but is less clear than coordination, is the number of food pellets they receive. The iSH payoff matrix provides asymmetric outcomes as anti-coordinated choices to the cooperate arm leads to an outcome of 0 (S), whereas an anti-coordinated defect choice (T) leads to the same outcome as a coordinated one (P). On each trial, the maximum outcome available to an agent animal is achieved by coordinating with the stooge. Figure 2.2.2b shows the mean proportion of the maximum outcomes available that agent rats received. Despite coordinating on 67% of sessions, animals received on average 0.8 of available outcomes. The payoff matrix in this task provided a constant reward for defection regardless of whether the agent animal defected in coordination with the stooge. Anti-coordinated cooperation trials (S), on the other hand, yield an outcome of zero. This asymmetry, in the context of pseudo random stooge, meant that anti-coordinated defection would only have led to a moderate decrease in overall outcomes. Notably, agent rats did actually not receive greater outcome than consistent defection, though did better than chance (t_{74}=16.7753 $p>0.0001$) yet continued to coordinate equally regardless of whether the stooge had cooperated or defected.

A key requirement in the Stag Hunt is the capacity and motivation to coordinate, which rats have demonstrated a consistent, yet moderate,

tendency to do so. In the prior chapter, we discussed in depth the nature of Nash equilibria, a collective choice from which neither animal can benefit by changing their choice. To determine the capacity agent rats have, to perseverate in an equilibrium, we exposed 4 naïve agent rats to stooges that cooperated on every trial. This all cooperate stooge is essentially an 'ideal opponent'. We then measured the capacity of the agent rat to coordinate, and therefore cooperate under these ideal conditions. We elected to employ an all cooperate stooge as the contrast between the two possible outcomes for an all cooperate stooge (3 vs. 2 pellets) is lower than that provided by an all defect stooge (2 vs. 0 pellets). This provides a somewhat more difficult decision to the agent and thus, is a better measure of their capacity to perseverate. We are therefore examining the capacity of agent rats to perseverate in behaviour consistent with the payoff dominant equilibria, coordinated cooperation against an entirely cooperative opponent.

2.2.3 Interim Summary

The purpose of these experiments was to determine if rats produce behaviour that is consistent with an understanding of the economic terms of the Stag Hunt by measuring their behaviour whilst controlling the behaviour of a stooge animal. In the first case, we employed a pseudo random stooge. A pseudo-random distribution of choices was used as this presents the agent animals with something very like a series of single games, as their opponent's choices are not influenced by the history of trials. Our interest was not to determine if agent animals could understand the strategy of the stooge animal. Rather we wanted to isolate the simple contingency between the stooge's position, the agent's choice and the

outcome, and determine if the agent rat behaved in a manner consistent with the inculcation of this association. Here agents demonstrate the appropriate behavior greater than chance, engaging in consistent but noisy coordination.

Next we asked if animals could produce behaviour consistent with the payoff dominant equilibria, mutual cooperation, by deploying an all cooperate stooge. This is done to establish and quantify a capacity to perseverate in coordinated cooperation, under ideal conditions. There is a difference of a single pellet between the R outcome yielded by cooperating with the stooge, compared to 2 pellets for the T outcome for defecting against it. Crucially this tendency does not exceed chance until the second session and further sessions until the behaviour stabilizes at its high levels. In the context of two freely behaving animals, this data suggests that it is in principle possible for rats to engage in behaviour consistent with ideal equilibrium predictions.

In each experiment agents have displayed behaviour consistent with an understanding of the economic terms of the Stag Hunt. However in each case behaviour could also be explained by a simple preference for following a conspecific. In the all cooperate stooge case, it could even be explained by merely developing a side preference and ignoring the stooge animal.

2.3.1 Anti Coordination Experiments

The prior experiments show behaviour consistent with an understanding of the economic terms of the Stag Hunt. However, agent

rats may have been interested in maintaining proximity to the stooge rat rather that in increasing the magnitude of received outcomes. To distinguish between these two possibilities, we implemented a modified 'anti coordination' version of the iSH payoff matrix (AC payoff) (table 1.3). This outcome structure reverses the outcomes for coordinated and anti-coordinated choices so that coordinated cooperation (R) yields the lowest outcome (0) and anti-coordinated cooperation (S) yields the highest outcome (3). This dictates that while defection will still yield a constant outcome of two pellets, the optimal choice for an agent rat on any given trial is now to choose to go to the opposite side from the stooge.

	Cooperate $_{agent}$	***Defect*** $_{agent}$
Cooperate $_{stooge}$	1,***0 (R)***	1,***2 (T)***
Defect $_{stooge}$	1,***3 (S)***	1,***2 (P)***

We will now proceed to describe both, how animals that have previously experienced the standard iSH payoff structure and naïve rats behave in the context of this AC-iSH payoff structure.

2.3.2 Anti-Coordination Experiment Results

To first determine whether under the standard payoff matrix agents rats were simply following the pseudo random stooge to maintain proximity to conspecific under the coordination payoff matrix or whether they were doing so to increase reward, we transitioned four dyads from the standard coordination condition to a novel anti-coordination condition. We continued to fix the stooge behaviour as pseudo-random. Here there are

two predictions that support outcome driven behaviour, first that agents will now show an increase in anti-coordination, the optimal behaviour under these conditions. Second they may adopt a sub-optimal behaviour in defection choices as this represents a safe moderate outcome.

Upon transitioning to the AC payoff condition, 4 dyads that have previously engaged in consistent coordination rapidly reduced this tendency to follow the stooge animal (fig 2.3.1(a)) showing significantly less coordination in the third session than the final session of the preceding condition (paired sample t_3=-2.938 ±0.120 $p<0.05$). Under the AC payoff matrix, rather than starting to anti-coordinate, agents rats opted instead to increase defection choices (fig 2.3.1(b)) showing significantly higher defection trials in the final session under that anti coordination payoff matrix than preceding it (paired t-test t_3=4.351 ±0.351 $p<0.01$).

To determine if this change from coordination to increased defection led to increased payoffs, we examined the proportion of maximum outcomes achieved in the session prior to the change in payoff in the session following and final session. Directly following transition to the AC structure where animals maintain a similar level of coordination, there is a sharp decrease in received outcomes (paired t-test t_3=10.78 $p<0.01$) (fig 2.3.1(c)). By the third session when agents are defecting much more often, we see that they have indeed increased their received proportion of the maximum reward available (paired t-test t_3=8.256 $p<0.01$). This indicates that while agents did not adopt the more lucrative strategy of anti-coordination, they did rapidly change their behaviour in response to changes in the contingency between their choice, the stooge's position and outcomes, which suggests that their tendency to coordinate

with a stooge animal under the standard payoffs is at least in part motivated by outcomes rather than simply following the stooge animal.

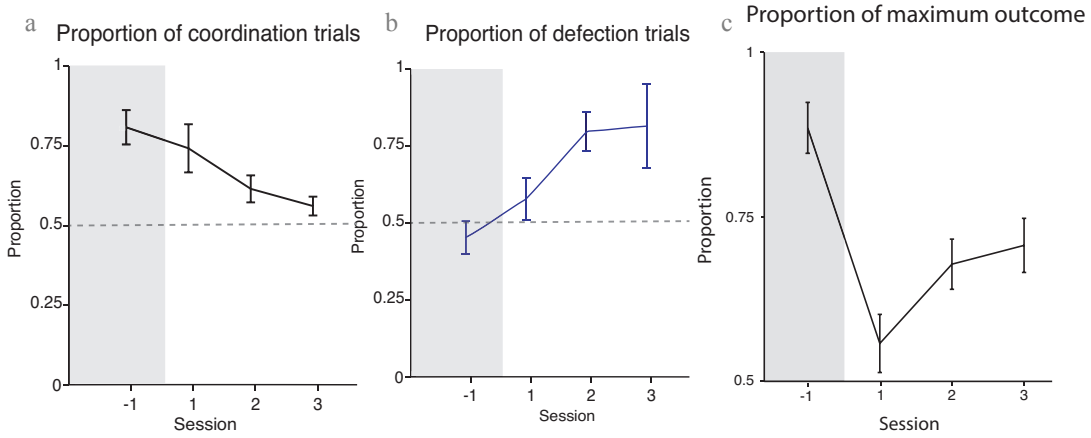


Figure 2.3.1 |Transition to anti-coordination iSH increases unilateral defection
Proportion of coordination (a), defection (b), and maximum outcome received (c) trials for four stooge agent pairs. Grey area indicates final session under standard iSH outcome structure while in sessions 1-3 agents continued under the anti-coordination payoff matrix. Error bars are SEM.

It is possible that this increased tendency to defect rather than anti-coordinate is influenced by their prior experience under the standard payoff contingencies or a failure to learn the novel ones within three sessions. We therefore conducted the same experiment but this time with eight naïve agents. Naïve agent animals competing against a pseudo random stooge with the AC payoff continue to display a similar constellation of behaviours. Here agent rats coordinated at a level greater than chance on average (t_{38}=7.110 $p<0.001$).

These data clearly show that agent animals are not engaging in anti-coordination, nor are they engaging in similar levels of coordination as in the prior condition. Looking at the distribution of outcomes (fig 2,3,2b) by applying a repeated measures ANOVA revealed significant effect of outcome upon proportion of choice ($f_{3,107.9185}$=0.6745,$p<0.001$).

Post hoc analysis using Tukey Cramer HSD demonstrated that animals are predominantly defecting as P (M=4.072 ±0.019) outcomes were greatest followed by T (M=0.3210 ±0.021), both of which were greater than R (M=0.1679 ±0.022) and S (M=0.1040 ±0.017) outcomes. Surprisingly, agent rats in this condition tend to cooperate more in coordination with the stooge, yielding zero pellets (R), than when the stooge defects, which would yield the highest reward of three pellets. This may be because they value the proximity of the stooge rat sufficiently to occasionally deviate from their tendency to defect. However this behaviour was quite rare as both R and S outcomes occurred less than chance (R: t_{38}=-7.101, *p<0.001,* S: t_{38}=-13.582). It is also important to note that as the proportion of cooperation and defection actions made by the stooge animal was pseudo random. As a result, agent animals engaging predominantly in defection will achieve chance levels of coordination purely by the fact the stooge produces defect actions on approximately half of all trials.

Despite predominantly defecting, the proportion of maximum reward is much closer to what would be achieved if agent animals had pursued an entirely random strategy and is both significantly less than the proportion of maximum reward achieved by agents in the normal condition (paired t_{112}= *p<0.001*) and that of an all defect strategy (t_{38}=-13.585 *p>0.001*) but much greater than they would have achieved by consistent coordination (t_{38}=-19.42 *p<0.001*). Importantly, following a perfect coordinate strategy would yield only 0.4 of the available reward, considerably lower (t_{38}=-68.72 *p<0.001*) than that achieved by agents in this condition.

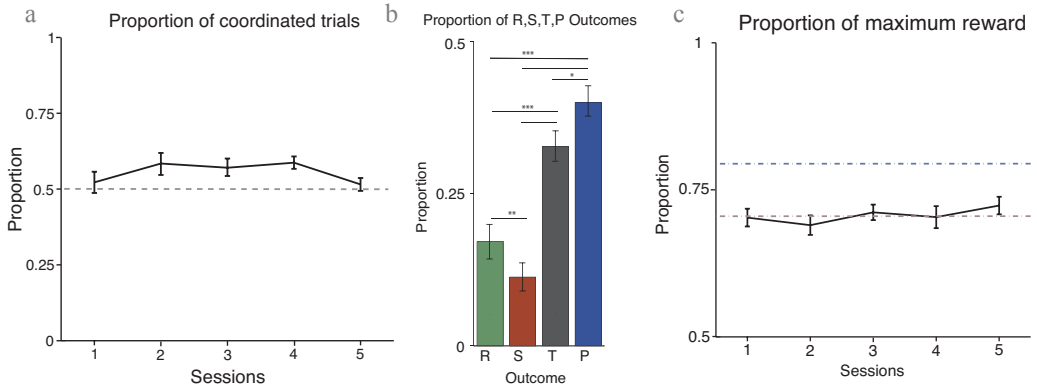


Figure 2.3.2 |Proportions of coordinated trials and outcomes.
Mean proportion of: coordinated trials by session (a), outcomes (R:green,S:red,T:grey,P:blue) across all sessions (b), (c) shows the mean proportion of maximum reward (provided by complete coordination with the stooge) received over sessions. Grey dashed lines indicate chance, blue, proportion of maximum outcome received by 100% defection, red, proportion of maximum outcome achieved by random behaviour.

To more clearly illustrate the different patterns of behaviour of agent animals facing a pseudo random stooge under the normal and anti-coordinate payoff, we will now compare coordination and defection from each condition. Coordination is a much larger factor in the behaviour of animals exposed to the standard rather than AC payoff (paired t test t_{112}=5.5167 *p<0.001*). Defection, on the other hand, is no different from chance in the standard payoff, but considerably greater in the AC payoff condition (paired t test t_{112}=6.327 *p<0.001*).

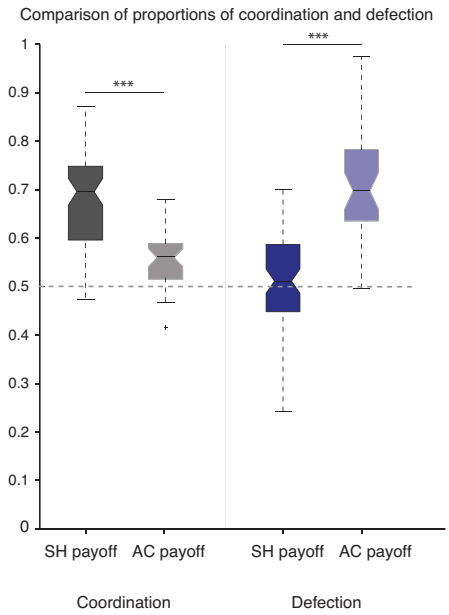


Figure 2.3.3 | Proportion of coordination and defection in each condition. Notched box plots showing the mean proportions of coordination (grey) and defection (blue) in the standard (dark) and AC (light) payoff conditions.

2.3.3 Interim Summary

These two anti-coordination experiments were conducted to determine whether the behaviour of agent rats was best explained by a preference to maintain proximity to the stooge animals or driven by outcomes. The above data suggests that both factors are important in driving the choices of agent animals, however neither are sufficient to explain it entirely.

Agent rats rapidly changed their pattern of choices from one best described as noisy coordination to a noisy all defect behaviour within two sessions following transition from the standard coordination payoff to one that encourages anti-coordination. This change to noisy all defect

behaviour also had the effect of increasing outcomes beyond what would have been received had they continued to coordinate. This clearly states that agent rats are motivated by received outcomes and while in both conditions they displayed some tendency to follow the stooge this was not the sole driver of behaviour in these experiments.

Naïve rats displayed a very similar patter of behaviour with defection rather than coordination being the primary best descriptor of their behaviour. Coordinated cooperate trials occurred much less than chance would predict, but, given high rates of defection, even a very low level of such cooperation will mandate that coordination will occur greater than chance.

2.4 Discussion.

The data from the above experiments shows that rats have a capacity to associate their choice, the stooge's position and outcome. Further we excluded the possibility that this behaviour was purely a result of the agent following the stooge by observing a change to safe noisy defect strategy under the AC payoff structure. Agent rats displayed a capacity to reliably perseverate with an all cooperate stooge, however this behaviour took two sessions to be clearly established. This indicated that rats are capable of producing behaviour consistent with the payoff dominant equilibria.

In the first experiment where agent rats played against a pseudo random stooge under the standard iSH payoff, they displayed a consistent preference for coordination. However, they did so only moderately. There may be two methodological explanations for why higher rates of

coordination were not consistently observed. First is that the animals used in this condition were pooled from several experiments. This assay was set up and moved several times once from one institute to another and each time this was done, the coordination experiment was re-run as a validation step before conducting further experiments. Some groups of animals performed better than others. The final cohort of animals, those that were also used in the transition to the AC payoff, coordinated at closer to 75% notably higher than the population as a whole. The assay was not moved from this location and conditions were kept almost identical for the experiments in the following chapter.

A further explanation for the moderate levels of coordination seen here relates to the structure of the assay itself. In these experiments, the start box and choice area were not separated, and the choice doors only opened following the stooge penetrating the IR beam at the end of the accessible arm. This meant that agent animals were often interacting with one or the other door for as long as a minute, before the stooge animal returned and then moved to the accessible arm. Agent rats often attempted to access one arm, then the other during this period. This meant that on some trials they were interacting with the door that would lead to an anti-coordinate choice when the opportunity to make their choice became available. This may have added some additional noise to their choice behaviour. In light of this, we added an extra door restricting access from the start area to the choice for the experiments shown in the next chapter.

Another possibility here, that is difficult to control for, is that while the stooge's reward was held constant, and its choices determined, any preference the stooge may have had for one or other arm on a given trial

was not within our control. We have not attempted to quantify such a preference. However, anecdotally, stooge rats did on some trials struggle with the door on the non-accessible arm. This may have affected the subsequent choice of the agent animal.

Despite this moderate level of coordination and these potential explanations not withstanding, agent rats in this condition did coordinate in a manner that demonstrates an understanding of the economic terms of the game. Like Viana *et al's* prior work, it is not possible in this case to state whether the stooge animal served as merely a cue, or whether other social learning strategies were being implemented. The purpose here was to show that animals are capable of performing this task so that it can be studied in the more complex case where both animals can choose freely.

One concern the reader may have is that the animals failed to anti-coordinate under the AC payoff conditions. From the data above, we cannot distinguish between this being a failure to accurately associate the contingency between the stooge choices and outcomes under the AC payoff conditions or whether this was the result of a conflict between a preference for proximity and a preference for higher outcomes. Tervo *et al*'s findings indicated when animals were not able to counter-predict the behaviour of their opponent, they defaulted to safe stochastic strategy. In the iSH task, the safer option, rather than behave stochastically, is to increase the frequency of defection choices, which is what both naïve rats and those that had previously been exposed to the standard payoff structure. While here there is no stooge strategy to predict in this case, a failure to accurately associate the stooge's behaviour with received outcomes would represent a similar result to the animal. It would represent

an inability to predict rewards and may result in a similar preference for the safer option.

A possible explanation for the animal's apparent capacity to reliably coordinate yet not anti coordinate with a stooge animal in each condition may be related to local or stimulus enhancement. Under the standard condition, animals coordinate greater than chance from the first session. The learned associations between stooge choice and outcomes may well be supported by such a mechanism. In the same way as described in the introduction, agent animals orient their behaviour towards where the stooge is, therefore following, which in turn supports more instrumental acquisition of the outcome contingencies over trials. While this may support acquisition of the coordination payoff, it is quite possible that this may inhibit learning this acquisition in the AC payoff condition. Here, to achieve the maximum outcome, the agent rat must associate the side *not chosen* by the stooge with the highest outcome. This use of social information in this manner has not been demonstrated in the rat. Masuda *et al* have demonstrated transmission of aversion behaviour however, in this example, animals that had prior experience of negative outcomes were more likely to avoid an area that a conspecific indicated aversion to. This however was still mediated by maintaining proximity to a demonstrator in the 'non aversive' area. While again we cannot confirm whether local or stimulus enhancement played a role in determining choice behaviour in each of these conditions, it does provide an ergonomic explanation for it.

What this data does suggest is that in freely behaving animals, both 'noisy coordination' and 'noisy defection' are two particular patterns of behaviour we can expect to see in the freely behaving case. Further as it is

unclear precisely why animals defected under the AC payoff, this reduces the effectiveness of this as a possible experimental condition to explore in the freely behaving case.

As we discussed in the prior chapter, Nash equilibria are an important aspect of the Stag Hunt task. By employing an all cooperate stooge, we explored the capacity of rats to perseverate in cooperation and thus behave in a manner consistent with that equilibrium. Rats did display a clear capacity to engage in such behaviour. However, unlike coordination with a pseudo random stooge under the same payoff conditions, this took several sessions to establish. This indicates that despite the question being asked of the agent being somewhat simpler as they only have to choose between two options (R or T) across all trials rather than the R, S, T, P against a random stooge. This may indicate that agent animals have a preference for alternation over perseverative behaviour. This does indicate that animals have the capacity to perseverate in cooperation, but even in the context of an all cooperate stooge, takes a considerable number of instances to be established. This suggests that in the freely behaving case, while establishing similar high rates of cooperative perseverance is in principle possible, it is unlikely to happen quickly and will require considerable tolerance in at least one member of dyad to low S outcomes.

Having established that rats do display the requirements to engage in the iSH task, we will now proceed to the next chapter where we will examine the behaviour of freely behaving animals.

Chapter 3

Simultaneous Choice

"Where do all these highways go, now that we are free"

Leonard Cohen, stories of the street.

3.0.1 Introduction

In the prior chapter, via the use of a stooge animal, it was demonstrated that rats have the capacity to engage in the iSH task. Demonstrator-observer paradigms, like the one just employed in the prior chapter, are ubiquitous in the study of social behaviour. By controlling, whether through conditioning, training, or direct manipulation, the behaviour of one animal and orchestrating interactions with another, one gains a powerful tool to examine the sensitivities and responses animals have to social information. Such a tool has allowed experimenters to hone in on particular cognitive elements that support or inhibit the likelihood of cooperation in game theoretical contexts.

An elegant demonstration of the experimental power of this tool was conducted by Stephens *et al* (2002) [21]who employed stooge animals to distinguish between two competing hypotheses for the lack of observed cooperation in the iPD. One conjecture was that animals lacked sufficient strategic sophistication to respond reciprocally, the second was that steep temporal discounting decreased the perceived value of the delayed, but on average increased reward that can be received from reciprocal cooperation. Here, the behaviour of agent scrub jays was observed in interactions with stooges that consistently defected and those that reciprocated. This was performed in conditions where reward was received directly following each trial, or accumulated visibly but was only accessible after a delay. This accumulation condition was implemented to mitigate the effects of temporal discounting. Here they demonstrated that there was an increased propensity to cooperate with a reciprocating stooge when rewards accumulate compared to when they were directly available. This was

found to be, in part, determined by an increase in what they termed 'forgiveness', a decrease in the likelihood of changing one's choice following a low S outcome in the accumulated condition. This 'forgiveness' effect was not present when agents played against an all defect stooge animal but was against one that reciprocated, showing that it was also modulated by strategy. This combination of experimental control of one animal and the timing of outcomes provided evidence that it is not a limitation in strategic understanding, but rather the presence of impulsivity that undermined reciprocal cooperation in the scrub jay[21]. The now much discussed work of Viana *et al* (2010) employed a similar approach, showing that rats were responsive to the strategy of the stooge, and to outcome and that this was also modulated by impulsivity[28].

This presents the question, given that controlling the behaviour of one animal and observing the behaviour of another that chooses freely is such a useful tool in examining the types of the social behaviour of animals, why attempt to study freely behaving animals at all? One argument in favor of such an approach is that it creates a more ecologically valid perspective upon cooperative behaviour. This however is not always the case. A compelling example to the contrary of this assertion can be seen in recent work exploring pro-social behaviour in rats. In this approach, rats displayed behaviours consistent with empathy. By observing rats' responses to the apparent distress of conspecifics that were trapped in a plastic restrainer by working to release them[71]. Further leveraging this paradigm, the authors demonstrated that acute anxiety states[72], and social experience[73] are important mediators of this behaviour. In this example, the rat that was imprisoned in the plastic container can be considered a demonstrator, its behaviour and the social information it

transmits being clearly modulated by its incarceration, while the observer would be its potential liberator. This particular experimental context does seem quite representative of the type of real world social encounter rats may actually have and it is not obvious how one might engineer a more ecologically valid case by not explicitly controlling the context in a similar manner. The same could be said for work examining the social transmission of fear mentioned in the prior chapter. By conditioning a demonstrator, one is able to control the timing and type of displayed fear responses in a manner that as a social interaction seems to carry meaningful ecological validity[62]. Indeed both these cases may be closer reproductions of the types of social interactions that rats engage in than the type of socio-economic interactions we are attempting to examine here.

There is a case to be made in that, in the context of a social dilemma game, having freely behaving agents rather than a stooge animal may indeed create a context that is closer to real world interactions in an important way. Game theory has two particularly defining assumptions, it formalizes the interdependence of interacting agents and it assumes those agents are rational utility maximizers[49,5]. However it is now well established that both human and non human animals very rarely, if at all, behave in a manner consistent with such rational expectations[20,38,40]. Stooge animals that embody idealized strategies, in contrast, do behave in a manner that is entirely born out of such rational expectations. For example, in the case presented by Viana *et al* (2011), reciprocal stooges were used whose behaviour was determined by, and did not deviate from, a perfect tit for tat strategy. Both, the particular nature and the precision of such strategies, have not been conclusively demonstrated in animals[5]. Similarly ideal pseudo random, all cooperate and all defect stooges may

simply not be strategies one is likely to see expressed in freely behaving animals. While animals facing stooges that embody such strategies may respond in a similar manner, this still does not provide evidence that such behaviour would emerge from an interaction when both can choose freely. This is not a criticism of the stooge-agent methodology, but rather highlights that the approach asks a different question than using freely behaving animals. For example, observing an animal cooperating with a reciprocal stooge asks whether that animal is *able* to engage in reciprocity, while examining freely behaving agents asks whether they *do* engage in such interactions. The same is true for the finding in the previous chapter that agent animals show behaviour consistent with the payoff dominant equilibrium in the iSH when paired with an all cooperate stooge does not mean that a freely behaving dyad will also establish the same behaviour.

The use of freely behaving animals, as a result, provides an opportunity to examine the manner in which animals establish cooperation, should it emerge, and the strategies they adopt to do so in a social dilemma task. However, this comes at the cost of losing some practical experimental control as one can no longer control the strategy of one player. And, as one is now examining the behaviour of two animals and their interactions rather than one, the complexity of analysis necessarily increases. It is perhaps for this reason that there are only very few examples of freely behaving social dilemma tasks in the rat.

A recent paper has examined the behaviour of freely behaving rat dyads that were visually separated but had access to auditory information in a weak iPD task. A weak Prisoners Dilemma is a variation of the classic game where the cost of unilateral cooperation is relaxed by making S and

P outcomes the same. Here, cooperation in one condition was considered to be the withholding of a lever press, and in a separate condition, with a different cohort of animals, this response association was reversed. In this case, they observed greater levels of mutual cooperation than chance when animals had to press the lever, this increased markedly following the fourth session indicating that learning was important for establishing cooperation. Cooperation did not exceed chance when animals were required to withhold a lever press to indicate cooperation. This serves to demonstrate that cooperation can emerge in weak iPD task with freely behaving animals, but also how particular idiosyncrasies within experimental design can have an important influence upon rates of cooperation.

A further, and slightly older, example, one that employed a similar dual T-maze design to the one we will use here, examined the behaviour of freely behaving dyads as they engage in a classical iPD task. In this case, they found that dyads displayed a low level of cooperation and when a barrier was placed between each T-maze blocking visual information, this level of cooperation decreased further. Perhaps disappointed by the absence of notable cooperation in their task, the authors pointed to the following anecdote as their more important finding. I quote directly from the paper below.

"The most "interesting" behaviors we observed are not reflected in any of the tables or graphs but rather must be classified as anecdotal. For instance, on Days 1 to 10, with the Plexiglas barrier in place so the animals could observe each others' responses, it was observed that the animals would rush from the start box and meet face to face at the

Plexiglas. There they would undertake an interesting "dance" whereby they appeared to feint left- and right-turning responses in what appeared to be an attempt to influence the other animal into making a desired response. Often 15 to 20 sec. of this mutual feinting behavior would occur before one animal would finally respond."

While this observation is anecdotal, and admittedly risks a degree of anthropomorphism, it does suggest that the interaction between the animals as they make choices is important. It also foreshadows the more conservative, but quantified finding of Marquez *et al* that rats are sensitive to the signaled preferences of conspecifics in a similar dual T-maze social assay[56]. What this observation does do is capture a subtle but important benefit of examining social dilemmas with freely behaving animals. It describes the sort of sensitivity to the ongoing flow of social information that seems a potentially important means by which animals may engage in collaborative actions and the types of behaviour that seem quite difficult to faithfully replicate using a stooge animal.

Due to the paucity of other examples of social dilemma tasks in freely behaving rat dyads, and perhaps in part to the popularity of the Prisoners Dilemma, there are no similar examples that have employed the Stag Hunt formalism. There are, however, examples where the central component of the Stag Hunt game, coordination, has been examined in freely behaving rats.

Schuster *et al* (2001) conducted an impressive and extensive examination of the behaviour in rats as they engage in a coordinated shuttling task, where animals had to alternate, back and forth in concert, to

and from a reward point to receive food rewards. The design of this task has similarities to the Stag Hunt in that coordination is central, however there was no cost for 'cooperation'. Animals are rewarded if they coordinate, and are not if they do not. Here, they compared coordination in contexts were animals were paired in the same assay, separated, but in yoked assays, and with access to a non-social light cue that indicated reward availability in each case. This revealed, firstly, that animals learned to shuttle in a coordinated manner more quickly and effectively when they were paired than when they were separated. However, when a non social cue, a light, was used to guide behaviour they learned to coordinate regardless of whether they were paired or separated and did so more quickly than when paired with an animal but did not have the light cue[32]. This demonstrated that freely behaving animals could effectively use ongoing social information to coordinate their behaviour. It also indicates that non-social information can also be a preferred cue over social information. This is perhaps because the behaviour of the conspecific was a more complex cue than the light. More fine grained analysis of behaviour in the purely social case also indicated that coordination in this task was non-stereotyped in that there was considerable variation in which animal would lead and which would follow, there were also frequent interruptions for social interactions and self grooming. This was emphasized by comparing how fast animals learned to shuttle in coordination with a naïve animal after learning to do so in a social context versus an isolated one. Novel pairings with animals that had preciously learned to shuttle socially learned to coordinate faster than those with an animal that had previously learned to do so in isolation. This suggests that rats learn something about how to work together through social experience that cannot be explained by proficiency in shuttling alone.

A further and important finding from this set of experiments was that coordination appeared to have an effect upon appetitive behaviour. Animals consumed greater quantities of a separate reward following coordinated social sessions than ones where they shuttled alone, even though they had received equivalent amounts of reward during the task in each case. This suggests that the act of collaborating with another rat can have the effect of increasing reward directed behaviours in itself.

The Stag Hunt has been explored in other species. Duguid *et al* (2014) examined pairs of chimpanzee's and infants (4 years old on average) as they engaged in a dual 2AFC Stag Hunt. They did so in both high and low risk conditions and with either the ability to see each other as they chose, or not. Here, they found that in all conditions human children coordinated and cooperated in the vast majority of trials. Chimpanzees, on the other hand, engaged in coordinated cooperation when they could see each other, though lower than that in children, and lower still when visual information was removed. They also engaged in coordinated cooperation less in the high-risk condition, which is most analogous to the iSH payoffs we defined in chapter 2, than in the low risk condition. Here they also quantified the extent to which their participants visually tracked and communicated with their opponent. Neither the human children nor the chimpanzees demonstrated any evidence of communication prior to choice. Communication was observed following choices and in the case where visual information was available, both the infants and the chimps visually monitored their opponent prior to choice.

Each of these examples demonstrates that, while sacrificing the control of a stooge animal, examining coordination and cooperation in freely behaving animals is entirely possible and potentially provides a novel richness of data derived from the interaction between animals. The importance of these interactions in each case highlights a concern with following a standard or evolutionary game theoretical approach to cooperation. Each of these approaches is agnostic to whether a cooperative act is performed in a social or individual manner. Social Dilemma games in their pure form are expected to be performed in isolation, where each player acts and receives outcomes alone[5]. The examples above, however, indicate that there are social aspects of interactions that are both ignored by these approaches and press upon their expectations.

The examples above demonstrate that information preceding, during, and following choice can modulate incentives to cooperate. The example of coordination in rodents indicated that the process of collaboration in itself has components that require learning and that are separate from simply learning to maximize reward. It has also been demonstrated that the act of collaboration can change the value of those outcomes. These factors, of course are also present in examples that employ a stooge animal. Often social information is permitted and collaboration with a stooge animal may also have effects upon the value of received outcomes. Also cognitive constraints, such as temporal discounting, can cause agents to fall short of expectations of rational behaviour[7,21,28]. With freely behaving animals however these effects are emphasized as they now influence the behaviour of both players and, of course, the interactions between them.

This raises the question of how to best make use of game theoretic predictions upon explaining choices in an interactive social dilemma task, when it is likely these predictions will not capture all of the relevant variables that determine choice. Here we recommend an approach of cautious integration, using game theoretical predictions to specify optimal expected outcomes and related behaviour. At the same time, however, designing experiments in such a manner that can also quantify the role of factors known to lead to deviations from these behaviours.

One step to support this is to attempt to quantify the extent to which animals satisfy or fall short of these predictions in a more controlled context. This we began in the prior chapter by showing that animals are able to satisfy economic predictions of the Stag Hunt by coordinating with a pseudo random stooge to increase their reward, and to consistently cooperate with an all cooperate stooge as predicted by the payoff dominant Nash equilibrium. Here we have also demonstrated that rats have limitations in their proficiency in each, coordinating only moderately to increase reward, and requiring learning to robustly persist in the payoff dominant equilibria. Further it shows that rats have a greater capacity to increase their outcome when it involves coordinating with a conspecific, than when it requires that they anti-coordinate. As Woods (2016) demonstrated in their analysis of freely behaving animals engaging in a weak iPD task, the manner in which a particular experimental approach instantiates the action of cooperation can also influence the likelihood of cooperating. This suggests that playing close attention to whether the experimental design in itself encourages stereotypical aspects of behaviour that may support or constrain their capacity to engage in economic choice behaviour. Of all of these constraints upon normative economic behaviour,

the social interactions of the animals are likely to be the most pervasive and simultaneously the most interesting. One way of exploring this is to manipulate the presence or absence of this information or its meaning and examine how such manipulations influence the decision making of freely behaving dyads.

In the following experiments, we will follow this approach and examine the behaviour of freely behaving and interacting rats as they engage in an iterated iSH. We will examine the role of social interactions by manipulating the amount of social information and its economic value.

3.0.1 Animal Subjects

The experiments were performed using 26 pairs of male non-litter mates of the outbred Sprague Dawley rat strain, from Charles River, Barcelona, Spain. All animals were housed in pairs under 12 h light/dark cycle. Experiments were conducted during the light period. Before starting the experiment, agent rats were habituated to the experimenter alone for one week and for approximately 10 days to the automated double T-maze assay and to the novel food used for the positive reinforcements in the iSH stooge task. Each experiment used naive cage-mate pairs (unless stated otherwise). All rats were kept at 95% or higher body weight and while not food deprived were given access to a limited number of 14 standard chow pellets after each session. They were given free access to water.

3.0.2 Training

Here (the training is identical to the prior chapter but repeated here for the reader's convenience) the sequence for agent rats and stooge rats

was the same. Firstly rats were handled for one week in the experimental room and introduced to the food rewards. Second they were gradually introduced to the iSH assay. The iSH assay is a complex environment with automated moving doors and pellet feeders that limit the movement of rats within the assay, in particular, the order and directionality of their choices. The assay, primarily due to these automated doors, is initially mildly anxiogenic to the rats, and thusly they are introduced to the maze with all doors open. Over the ten days of habituation to the environment, they are introduced to these doors and during this period, they also learn the location of pellets delivered by the pellet feeders. Once each rat moves comfortably around the maze with all doors active and receives a single pellet reward for each completed trial, displays no significant bias to either side, they are considered ready to start the experiment.

3.0.3 Statistical Analysis

All the analysis was performed in Matlab 2010 (The 7.10.0 Mathworks, Inc). Means and standard error of the mean will be reported. All data has checked for assumptions of normality unless otherwise mentioned. General linear modeling was performed in R.

3.1 iSH: Equal Rewards Condition

The iSH task is based upon the provision of outcomes to two interacting animals as a function of their combined choices. During this task, animals will have access to ongoing social information as they make these choices and while they receive their outcomes. This social information may have an important effect upon important aspects of choice dynamics, such as coordination, regardless of whether these choices change lead to increased rewards. Animals will also likely show behaviour

that is particular to the assay itself, which are not a function of economic outcomes, that each animal will be forced to account for in their interactions with their opponent, which will also be measured in this case.

To examine this we implemented the following Equal Rewards condition. Here each animal in a dyad had access to ongoing social information but rather than the outcome each received being a function of their combined choices, a constant reward of two pellets is given regardless of what each animal chooses. In the terms of iSH payoff matrix R, S, T and P will all equal 2 (see table 3.1). This allows us to generate an ad-hoc baseline for the purely social aspects of behaviour we are likely to encounter in the iSH assay. Our purpose here is to determine what are the tendencies for animals to engage in behaviour relevant to our analysis, primarily coordination, but also their propensity to develop side preference, when outcome is constant.

Table 3.1

	Cooperate	***Defect***
Cooperate	2,***2 (R,R)***	2,***2 (T, S)***
Defect	2,***2 (S,T)***	2,***2 (P,P)***

3.1.1 Equal Rewards Condition Results.

Six naïve dyads, trained in the manner detailed above, performed eight sessions under the 'Equal Rewards' conditions. Data from both, Coordination and Anti-Coordination stooge experiments suggested that dyads have a preference for proximity to a conspecific. In the former, coordination was reinforced, in the latter, coordination was maintained at a low level despite being penalized. Coordination has an important influence

upon the outcomes animals receive for cooperation in the iSH task. In that context, the greater the propensity for dyads to coordinate, the less likely either animal is to receive a 0 pellet S outcome rather than the 3 pellet R outcome. Here we quantified dyads' preference and capacity for coordination when both could choose freely, both had access to social information, but reward was held constant.

Firstly, we examined the proportion of trials where each rat in a dyad made the same choice. This 'simple' coordination captures the tendency for dyads to make collaborative choices, but will also include trials where each member of a dyad is simply acting upon a shared preference. To help disentangle more collaborative choosing from shared preference, we also calculated the conditional probability of both animals choosing the same side when one or other animal changed their choice from the prior trial. This we term 'coordinated switching', which more closely captures trial-by-trial synchronizing of choices by each member of a dyad.

Rats show a tendency to engage in both simple coordination and coordinated switching greater than chance (0.5) across all dyads (one-sample $t_{47}=5.9175$, $t_{47}=6.0541$ $p<0.001$). This moderate tendency to coordinate remains consistent over sessions (figure 3.1.1 b & c). Rats in this condition engage in each form of coordination to a similar extent indicating that coordinated trials are not a result of shared preferences for one or the other choice arm.

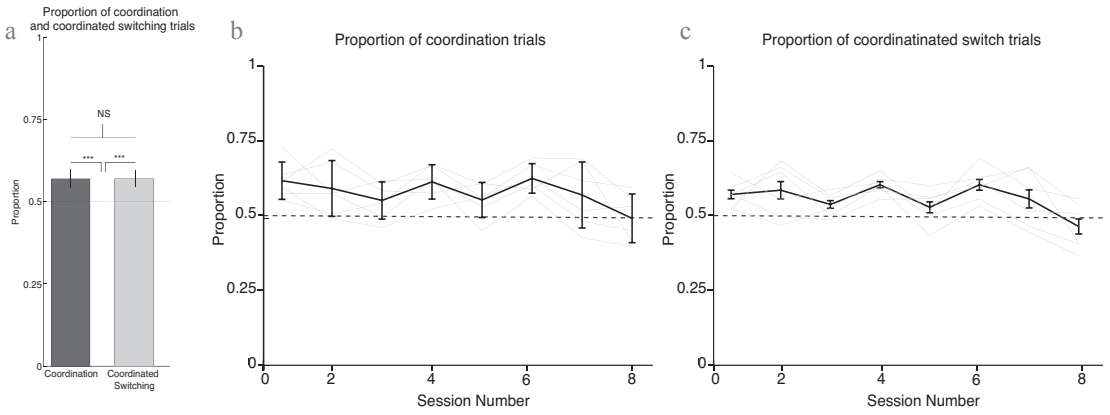


Figure 3.1.1 | Dyads show consistent moderate levels of coordination independent of outcomes. (a) Mean proportions of coordination, and coordinated switching across all pairs and sessions. Proportion of total trials where each agent makes the same choice by sessions (b) and the conditional probability of coordinating given either member of a dyad changed choice from the prior trial (c). Light grey lines indicate pairs, black line the mean, and error bars are SEM.

This establishes that rats have a moderate preference to coordinate their choices in the absence of economic incentives to do so. Animals here coordinated their choices in approximately 60% of trials, almost all of these coordinated trials occurred when one or the other animal changed their choice. This indicates that animals were changing their choices very frequently. To examine this in more detail, we generated probability mass density function over bout length, the number of uninterrupted choices to one arm or another (fig 3.1.2). Here we termed the choice arms cooperate and defect using the same convention as in other conditions for comparative purposes. This revealed that rats showed a strong tendency to alternate with the likelihood of single trial bouts of a given choice being much more likely than longer bouts of 2, 3 or more similar choices.

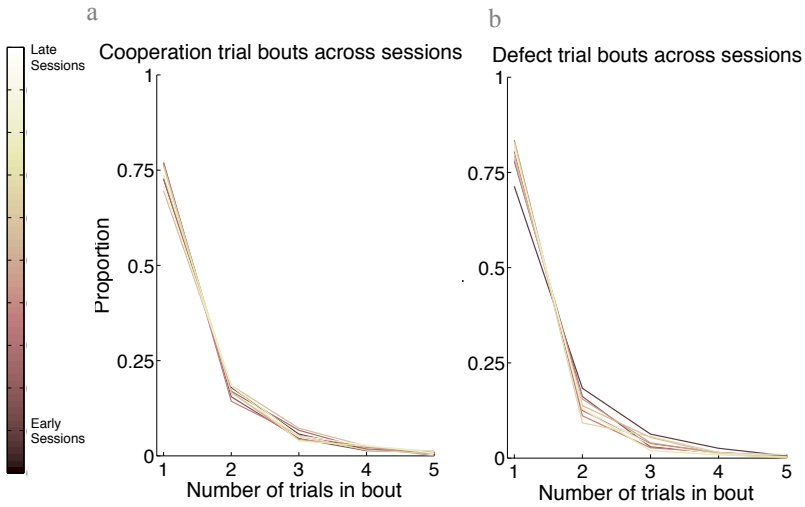


Figure 3.1.2 | Cooperation and defection bout length across conditions. Probability density functions for the number of continuous cooperation (a) and defection (b) trials across all rats, Color coded by session, darker to light for early to late sessions.

Having established a sort of ad-hoc baseline for coordination and for alternation behaviours, we then examined the extent to which individual rats within the dyad developed preferences for one choice arm or the other. As the outcomes available on the cooperate arms are equivalent, we calculated a preference index, which is simply the proportion of choices made to the left arm of the maze by either animal within a dyad.

Here we see that across all sessions and dyads there is no preference for one arm over the other. Within dyads however, we saw that six of twelve animals show a consistent preference for one or the other choice arm (one tailed t tests, all $p<0.05$ *(significance indicated in fig 3.1.3))*. Notably, within only one of the six dyads did each animal

demonstrate biases to opposing arms, which is further indication of moderate coordination.

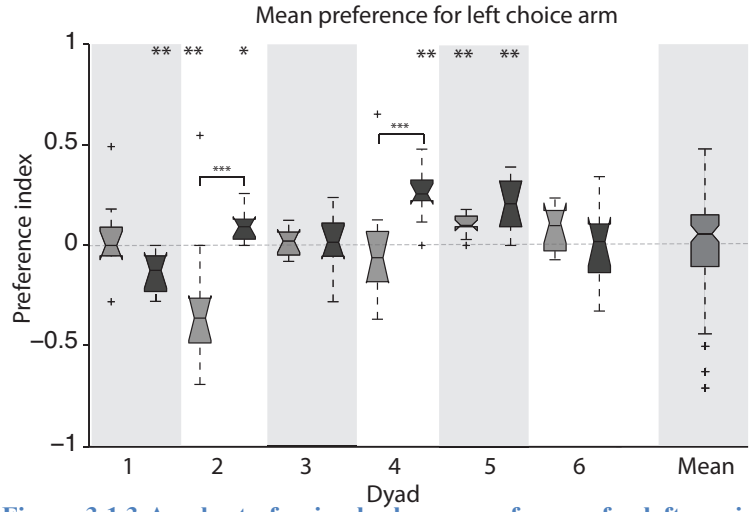


Figure 3.1.3 A subset of animals show a preference for left or right choice arm

Notched box plots indicating mean arm preference (proportion of left choice – proportion right choice) for each individual rat within a dyad across sessions. The population mean is shown on the right.

It is not unexpected that animals will develop a bias to choose one arm or the other. This data shows that even while doing so had no influence upon received outcomes, animals are equally likely to develop a spontaneous side preference as not to, and that across pairs there is no consistent preference for one or the other choice arm.

In the subsequent experiments, each choice arm will correspond to the iSH payoff matrix. One arm will be defined as cooperate and one as defect, whereas in this condition each choice arm is essentially identical.

For the purposes of comparison with other conditions, we assigned cooperation and defection to the choice arm for each dyad in the same manner as in subsequent experiments and analyzed the data on this basis. Here we see, as the above data predicts, that there is no significant difference between the proportion of mutual cooperation and defection trials (fig 3.1.4(a)). We see the same for the proportion of cooperation and defection trials with neither being greater than chance or necessarily different from each other (fig3.1.4(b)). Anti-coordinated outcomes where each rat receives either T or S outcomes are significantly lower than chance (t_{53}=-6.1807 $p<0.001$).

Figure 3.1.4 | Dyads show no difference in cooperation or defection. Coordinated cooperation; Reward trials (green) and coordinated defection, punishment trials (blue) across session(a), (b) mean proportion of cooperation and defection trials across all rats and sessions. Error bars are SEM.

Another important aspect of choice behaviour in the iSH task is which animal chooses first. Under the standard payoff conditions, choosing first and cooperating involves greater risk than choosing second, as whether a high R outcome or a low S outcome is received is contingent upon the other's choices. Choosing second, of course, removes this

uncertainty. Choices were recorded by an IR beam at the end of each choice arm.

Fig 3.1.5 shows the proportion of trials each rat in dyad chose first. In all but one dyad (multiple paired sample t-tests $p<0.001$ (significance indicated on fig 3.1.5)) there was one animal that chose first more often than its counterpart. As there is no particular cost here for one animal to choose before another, this may be due to individual or motivational differences within the dyad.

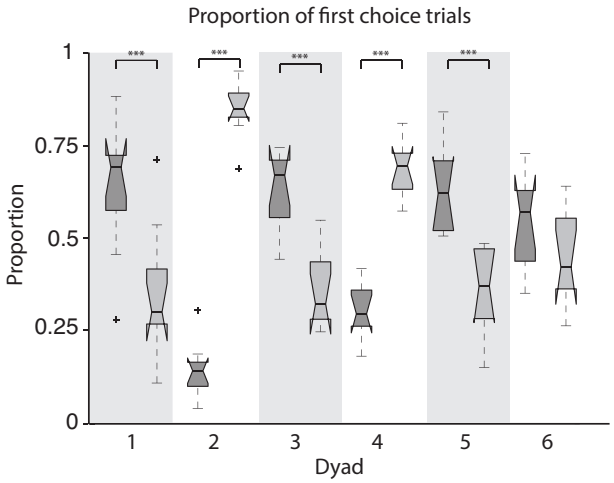


Figure 3.1.5 Dyads tend to have on animal that chooses first most frequently
Notched box plots indicating mean proportion of first choice trials for each rat in each dyad. Dark grey= rat one light grey= rat two.

Summary

The iSH task provides a means of examining how social information and contingent outcomes are integrated to guide choices. We conducted a simple version of the iSH task where all outcomes were

constant to determine spontaneous behaviours related either to social information or the experimental design itself that may influence the choices of animals in the subsequent conditions.

Primarily, rats demonstrate a consistent preference and a capacity for coordination, engaging in simple coordination and coordinated switching in approximately 60% of trials. Coordination is fundamentally important in supporting cooperation in the iSH task as unilateral cooperation leads to the unrewarded S outcome. Should dyads coordinate on 60% of cooperative trials in the standard condition, this would yield an outcome of 1.8 pellets on average, less than the outcome of 2 pellets for either coordinated or unilateral defection. This means that the levels of 'spontaneous' coordination alone would not be sufficient to entirely mitigate the risk of the cooperative choices in the iSH task and therefore not sufficient to support the establishment of a preference for cooperation over defection in the standard iSH task in itself.

An important finding above was that rats display a marked preference for alternation. Such a tendency for each animal to regularly change their choice reduces the likelihood of the simplest coordination strategy, choosing one side consistently. This makes both coordination, and thus lucrative coordinated cooperation a more complicated proposition. Conversely, it does increase the likelihood of animals experiencing all available outcomes rather than becoming rapidly fixed in one or other Nash equilibrium.

The data above shows that half of the population of animals developed a bias for one or other choice arm. This was seen to be present

in half of the animals in the Equal Reward condition. Such a preference in the context of moderate levels of coordination could support cooperation, however there was no consistent side preference across all dyads. This preference was only shared in one dyad and was not consistent across dyads. Therefore, if a similar side bias is present in the standard iSH, while it may support mutual cooperation in some dyads, it would be equally likely to support mutual defection.

3.2.1 iSH Standard Condition Results

Having examined the influence of a freely behaving conspecific on choice dynamics in the context of constant outcomes we now proceed to conduct the canonical iSH experiment. In this condition, we used the standard iSH payoff matrix described in chapter one and two and shown below.

	Cooperate	***Defect***
Cooperate	3,***3 (R,R)***	0,***2 (S ,T)***
Defect	2,***0 (T, S)***	2,***2 (P,P)***

Here, 10 dyads performed 14 sessions totaling 9395 trials, each dyad performed a mean of 670 trials (±7.91) and 67 trials (±1.94) per session, each lasting approximately 45 minutes. Fig 3.2.1 shows two representative raster plots illustrating the cooperate (green), defect (blue) and coordination (white) behaviour for two dyads. One important intuition to take from these plots is that dyads show a striking trial-by-trial variability in their choice behaviour. (video below)

Video 3.2
(To view video, download QR code reader from Google play store or Apple app-store and scan with phone camera)

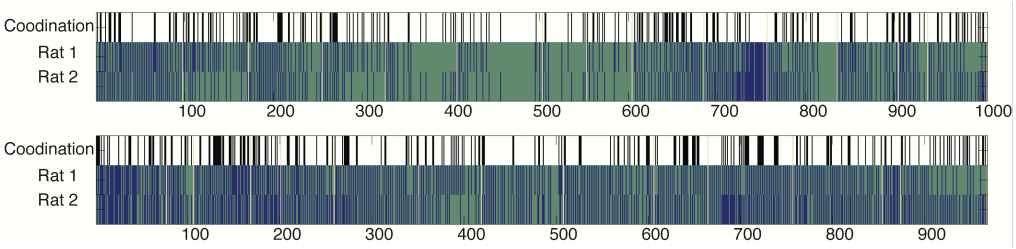


Figure 3.2.1 | Example Cooperate/Defect & Coordinate rasters of two dyads
First row shows coordination (white) and anti coordination (black) between cooperate (green) and defect (blue) choices made by rat 1(second row) and rat 2 (second row). Trials are shown on the x axis. Session boundaries are indicated in beige

In the standard iSH condition, coordination remains an important measure of performance. Rat dyads demonstrate high levels of simple coordination (M=0.71 ±0.09) and also coordinated switching (M=0.65, ±0.089) both of which were greater than chance (one tailed t tests t_{139}=26.2230, t_{139}=17,7042, p<0.001) (fig 3.2.3). Unlike the equal rewards condition, here the proportion of coordinated switch trials was greater than that of simple coordination (one tailed t_{139}=9.9407 $p<0.001$). Simple coordination can also represent the presence of shared preference for side, which is likely to explain, at least in part, the difference between coordinated switching and coordination trials.

Unpacking this data across sessions, we see a similar pattern to the 'all equal' condition with both simple coordination (M=0.68 +-0.785) and coordinated switching (M=0.63 ±0.097)) starting greater than chance

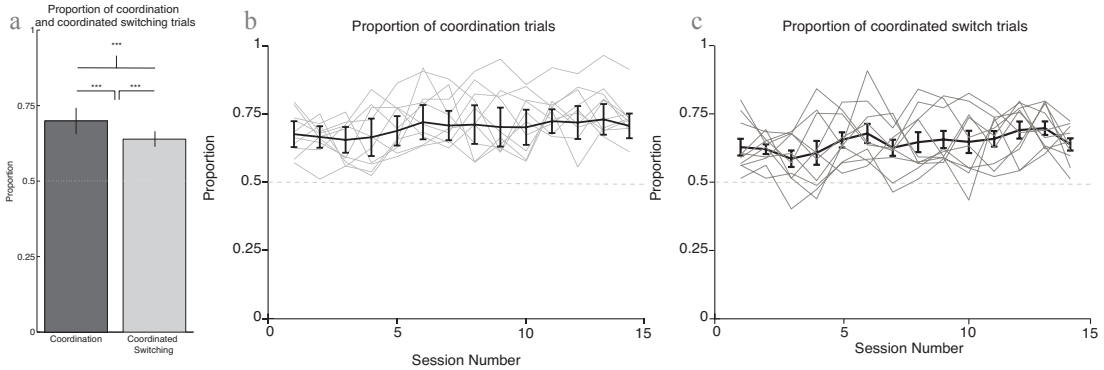


Figure 3.2.2 | Agent dyads robustly coordinate.
Mean proportions of coordination, and coordinated switching across all pairs and sessions (a) proportion of total trials were each agent makes the same choice across session(b) across all sessions, conditional probability of coordinating given either member of a dyad changed choice from the prior trial (c). Light grey lines indicate indivdual pairs, black line the mean, and error bars are SEM.

(t_9=7.280 *p<0.001*, t_9=4.2964, *p<0.01*). Simple regression analyses in each case showed no change over sessions (fig3.2.2), that each rat in a dyad is able to coordinate their choices with their partner satisfies the first expectation of iSH game, and is consistent with the behaviour of agent rats in the Stooge coordination experiments.

In this case, a large proportion of coordination trials consisted of those where one or the other animal changed their choice, indicating that, like the prior condition, animals maintained a preference for alternation. Again we see in fig. (3.2.4) that animals maintained a preference for alternation having been most likely to engage in bouts of only a single trial. However now, this changes over sessions with animals being increasingly likely to perseverate in cooperation and less likely to do so in defection.

Coordination trials consist entirely of either mutual cooperation (R) or mutual defection (P) trials. Mutual cooperation provides the greatest outcome (R) and thus is a measure of how well animals are collectively increasing their reward, and engaging in behaviour consistent with the payoff dominant equilibrium. In addition to the two outcomes available from coordinated choice, two outcomes result from uncoordinated choice S and T. Here we see across all sessions that dyads achieve R outcomes greater than chance (t_{139}=13.158 *p<0.001*). A repeated measures ANOVA revealed a significant effect of outcome upon proportion of choice ($F_{3,186.43}$=3.1672 *p<0.0001*), post hoc analysis by Tukey Cramer HSD indicated that the mean proportion of R outcomes (M=0.463, ±0.0126) was greater than that of P outcomes (M=0.2566 ±0.0097) and both were greater than anti-coordinated outcomes (M=0.141 ±0.089), (fig(3.2.3a)). As rats show a robust tendency to coordinate from the first session, we also examined the proportion of coordinated trials where R outcomes were achieved; this was also greater than chance (0.5) (t_{139}=7.2670 *p<0.001*).

Examining this across sessions, we see a different pattern than that observed for coordination. Here we see in the first session that mutual defection choices, consistent with the risk dominant equilibria, are greater than mutual cooperation choices (t_9=2.9421 *p>0.05*). A repeated measures one way ANOVA revealed a significant effect of session on mutual cooperation ($F_{13,126}$= 5.1798 *p<0.001*) with all sessions from the sixth session onwards being greater than the first. We see a similar though weaker decrease in mutual defection trials across sessions, ($F_{13,126}$=3.9591, *p<0.001*) (fig3.2.2b). This may indicate that dyads require a period of established coordination, which here is 5-6 sessions, before they are able

to regularly behave in a manner consistent with the payoff dominant equilibrium. However, following this 'learning' period, dyads are able to maintain a high proportion of choices consistent with that equilibrium.

Another important measure of the animals' propensity to cooperate is the simple proportion of cooperative choices made, regardless of whether or not those choices are coordinated. Across all sessions and animals, we see that animals cooperate more than chance (and necessarily defection) (one sample t test $t_{279}=10.124$, $p<0.0001$). This further confirms that animals are able to robustly cooperate in the iSH task.

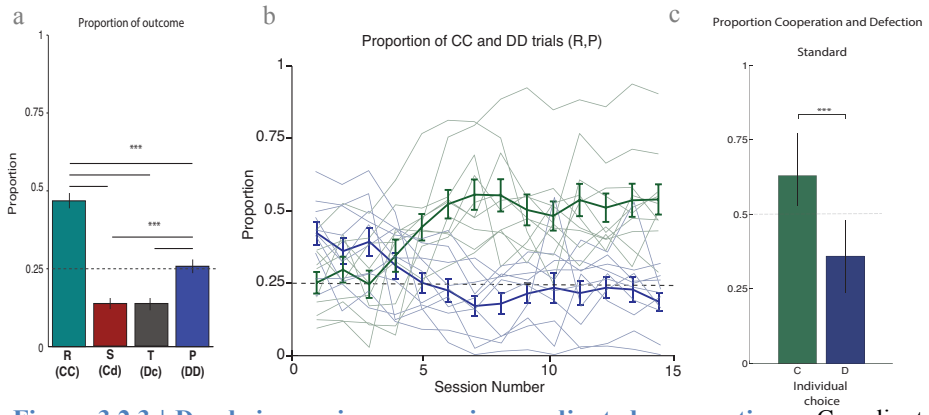


Figure 3.2.3 | Dyads increasing engage in coordinated cooperation. Coordinated cooperation; Reward trials (green) and coordinated defection, punishment trials (blue) across session(a) and all outcomes across sessions and pairs, reward (green), punishment (blue), anti-coordinated trials, temptation (black) and sucker (red) (b). (c) shows the mean proportion of cooperation and defection trials across all sessions. Error bars are SEM.

That rats engage in mutual cooperation at a rate approaching twice chance and do so greater than mutual defection is a central finding of this thesis. It demonstrates that not only can individual rats satisfy the expectations of the iSH task by coordinating but they are able to reliably

and collectively overcome the risk of the S outcome to repeatedly demonstrate behaviour consistent with the payoff dominant equilibrium.

The Nash equilibrium concept, in the context of Stag Hunt, predicts that animals will perseverate in their preferred equilibrium. Unlike the all equal condition, animals in this condition experienced the standard iSH payoff conditions and displayed a preference for the payoff dominant equilibrium suggesting that the previously observed alternation should rapidly decrease. Fig (3.2.4) shows probability mass density functions, (as in fig (3.1.2)) indicating the likelihood of differing lengths of continuous bouts of cooperation and defection. Here we see that, in line with an increased preference for cooperation, that longer uninterrupted bouts of cooperation are more likely over sessions, while bouts of defection become increasingly likely to consist of only a single trial. Notably however, animals still show a strong preference for alternation.

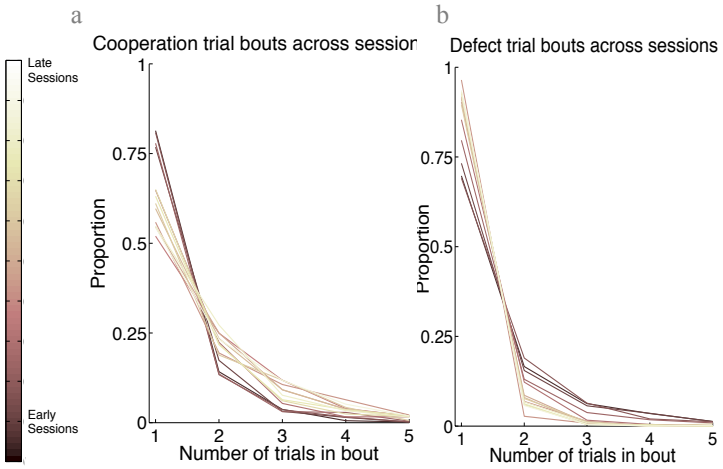


Figure 3.2.4 | Cooperation and defection bout length across conditions. Probability density functions for the number of continuous cooperation (a) and defection (b) trials across all rats, Color coded by session, darker to light for early to late sessions.

A further measure of performance, beyond coordination and mutual cooperation, though tightly linked to both, is the amount of rewards received. Fig 3.2.5 shows proportion of the maximum outcome (maximum outcome being achieved by engaging in 100% mutual cooperation)

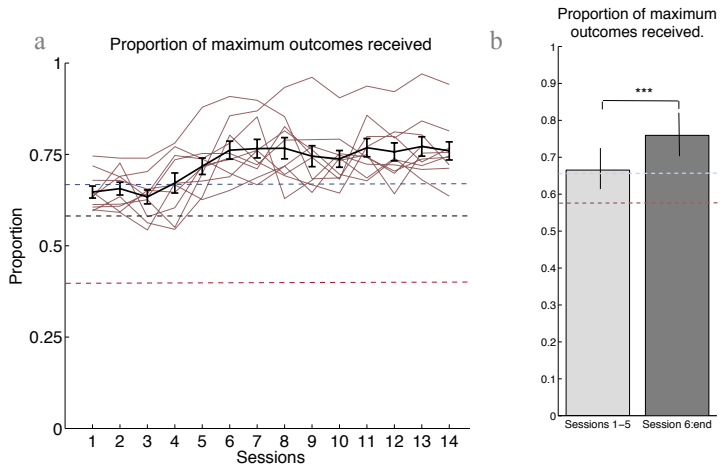


Figure 3.2.5 | Dyads effectively and collectively increase outcome.
Proportion of outcomes, the number of pellets received by the dyad as a proportion of the maximum reward available from an all cooperate strategy over sessions (a) and the mean across early (sessions 1:5) and late (sessions 6:14) (b). Dashed lines indicate the proportion of reward yielded from perfect 'all defect' (blue), random (black) or anti coordination (red) strategies. Black line is mean, brick lines are individual dyads and error bars SEM

achieved across *each dyad* and all sessions. Fig 3.2.5a describes, as expected, that the proportion of maximum outcome achieved increased in a similar manner to the proportion of mutual cooperation trials. If we now examine the mean outcome achieved in early 'learning' sessions (1-5) and later (6:14) sessions, there is a clear difference. In the early sessions, dyads achieve average outcomes that were greater than random behaviour (one sample t_{49}=8.2798 *p<0.001*) but indistinguishable from the outcomes that would be available from a 100% defection strategy. In later sessions however, dyads achieve greater outcomes than a 100% all defect strategy (one sample t_{89}=11.164 *p<0.001*), and the early 'learning' period sessions (paired sample t_{139}=7.0124 *p<0.001*).

In the iSH task, to cooperate *first* carries the most risk, as one cannot be sure that their opponent will reciprocate this choice. We therefore examined whether, within dyads, there was a tendency for one or the other animal to choose first. Here, interestingly, only two of the six dyads had one animal that consistently chose before its opponent across sessions (paired t tests, t_{13}=-3.142, t_{13}=5.5213, $p<0.0001$)) indicating that having one rat 'lead' is less consistent than in the 'equal rewards' condition.

As going first or second in the standard iSH has a potential influence upon received outcomes, we examined the proportion of outcomes, R, S, T and P for first choices and second choices (fig 3.2.5a). Across sessions and dyads, we observed no difference in the proportion of each outcome depending on whether an animal chose first or second (fig 3.2.5b). One might expect, for example, that sucker outcomes would be less frequent when choosing second than when choosing first, this however was not the case. This may be explained by a combination of high levels of coordination, which meant that regardless of which animal chose first, both animals made the same choice in the majority of trials. This suggests that the remaining trials may be accounted for by errors or exploration.

That fewer dyads have a stable 'leader' animal than in the 'all equal' condition may be due to the fact that in the standard condition there is a benefit to coordinating and a potential benefit to cooperation, making each choice a more complex proposition undermining the influence of individual or motivational differences upon choice order.

3.2.2 Standard iSH Condition Summary

The iSH task is a coordination task where the optimal choice for each rat is to coordinate its choices with its partner. This capacity for rats to coordinate their choices appears within the first session and is maintained throughout fourteen sessions indicating that each animal within a dyad is capable of performing behaviour consistent with the expectations of a Stag Hunt.

Perhaps the least demanding manner in which rats could solve the problem of coordinated action is simply to choose one side consistently and thereby cooperate or defect for the majority of trials. Both, the striking trial-by-trial variability, high levels of coordinated switching trials indicate, and an established tendency to alternate show that this is not the strategy being employed by rats in this task. The type of coordinated switching seen here requires that rats are either making use of ongoing social information as they make choices or have developed fairly accurate expectations of the other rats choice on a given trial, or some combination of the two. Notably, coordinated switching, while considerably greater than chance, is lower than coordination, implying that a meaningful proportion of anti-coordination trials occur either when these expectations are incorrect or ongoing social information is ignored or inaccurately interpreted by either rat.

The Stag Hunt is also a risk game. The distribution of outcomes defined in this iSH task highlights this aspect by making cooperation beneficial, if coordinated and costly, if chosen unilaterally. A cooperate choice on a given trial offers an increase in reward of 50% over defection

if coordinated, but a loss of 100% if unilateral, when compared against defection. Again, the simplest reward based strategy either animal could take would be to mostly or exclusively defect. Here we see that over sessions, dyads demonstrate a capacity mutually cooperate in approximately 50% of trials, twice that of chance. It is interesting that the majority of dyads do not increase mutual cooperation beyond this level, as there is no obvious constraint, having established a robust capacity to both coordinate and cooperate, that would limit them from cooperating more frequently.

The Stag Hunt is also referred to as a trust game, as one should cooperate only if they 'trust' that their opponent will do the same. Here we have seen that dyads begin with a higher rate of mutual defection than mutual cooperation in early sessions. The Stag Hunt predicts that if one is unsure of the choice of their opponent, then animals should defect and receive a guaranteed moderate reward. Data from the stooge experiments in Chapter 2 strongly indicates that animals are sensitive to changes in outcomes within a single session. This implies that this 'learning period' of five sessions may not be entirely accounted for by learning the contingencies of the game, but also upon animals learning to have confidence that their choice to cooperate will be reciprocated.

Interestingly, at the level of analysis approached thus far, choice order appears to have no obvious effect upon received outcomes. This is perhaps partly due to the fact that animals are given access to the choice arms simultaneously, and many of their choices are close to simultaneous, limiting the impact of choice order. There is also less of a tendency within a dyad for one or the other animal to choose first than in the 'equal

rewards' condition, which may be indicative of the increased complexity of trial-by-trial choices in the standard iSH task.

While prior stooge experiments indicated that agent rats are sensitive to outcomes, this does not exclude the possibility that high rates of mutual cooperation are supported by the development of a habitual preference in later sessions for the cooperate arm, rather than an active negotiation of choice.

3.3.1 Reversal condition

To determine the extent to which the dyads' cooperative behaviour might be supported by an established habitual preference for the cooperate arm, we conducted the following simple reversal experiment. Here we reversed the cooperate and defect arms for four dyads for a further seven sessions following completion of fourteen sessions of the standard condition.

Dyads that transitioned to the reversal task showed no significant change in coordination or coordinated switching from the last trial in the standard condition and a simple regression analysis reveals no significant change across sessions (fig 3.3.1a and b). Despite maintaining consistent levels of coordination and coordinated switching upon transitioning to the reversal condition, the underlying relationship between mutual cooperation and mutual defection trials changed considerably.

Figure 3.3.1(c) shows the proportion of mutual cooperation and mutual defection trials prior to the reversal, in the trial following and the

final session of the reversal. In the session preceding the transition to the reversal, there was a greater proportion of mutual cooperation trials than mutual defection (t_9=4.281, *p<0.01*). In the first session following the reversal, this difference is gone and is restored by the final session (paired sample t_3=2.9254 *p<0.05*). From the first session till the last of the reversal condition, simple regression analyses over sessions reveal an increase in mutual cooperation trials ($r^2_{0.666}$, $f_{11.4664}$, *p<0.001*) and decrease in mutual defection trials ($r^2_{0.7326}$, $f_{15.7514}$, *p<0.001*). These data again emphasize that individual rats in the iSH have a strong tendency to maintain coordination. However, following a change to the expected outcomes each dyad loses its preference for mutual cooperation over mutual defection. It then takes several sessions to re-establish similar levels of mutual cooperation on the reverse arms.

Further, if we examine the overall proportion of cooperative choices of individual rats between the standard condition and the reversal condition (fig 3.3.1(d)), they are equivalent and the proportion of cooperation trials in the reversal condition remain greater than chance (t_{55}=4.5865 *p<0.001*).

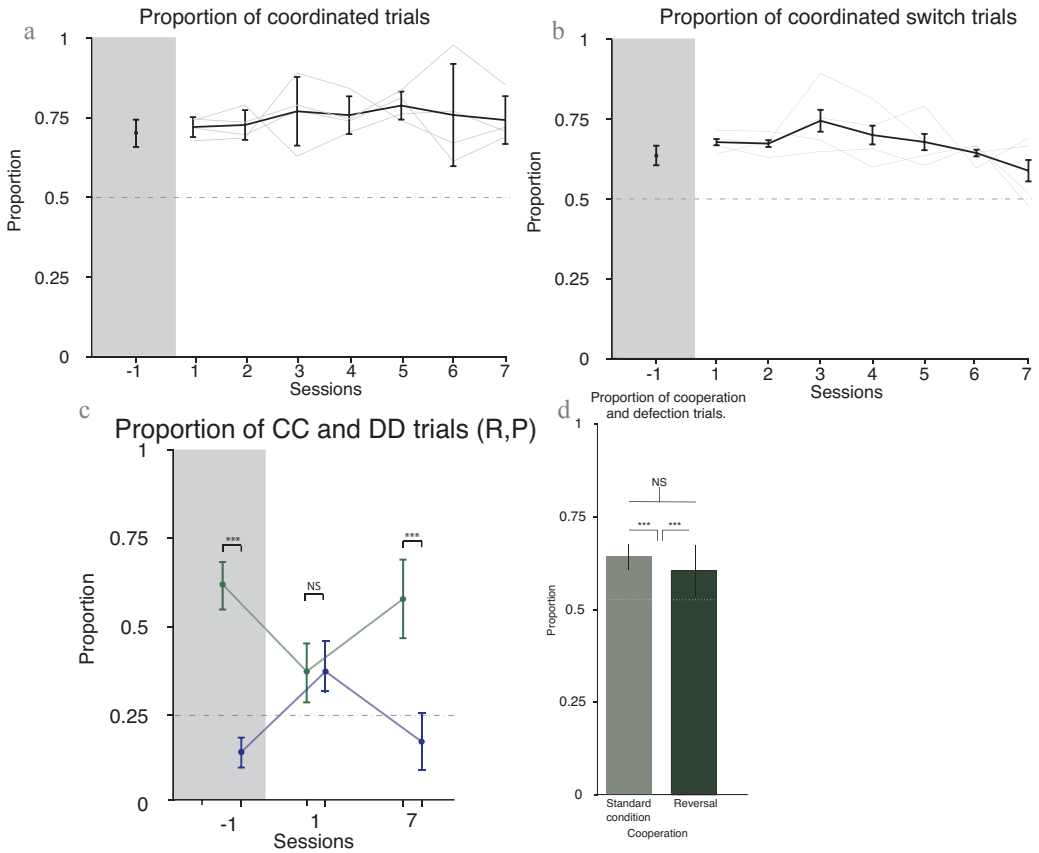


Figure 3.3.1 | Dyads maintain coordination and re-establish high mutual cooperation after reversal

Proportion of total simple coordination trials (a) and coordinated switching (b) before the reversal (in grey) and after. (c) shows the proportion of mutual cooperation (R) (green) and mutual defection (P) trials (blue) before (in grey), the first trial following and the last session of the reversal. (d) shows the mean proportion of cooperation trials for the standard condition and the reversal conditions across all pairs. Error bars are SEM.

3.3.2 Summary

The data from the reversal condition indicates that rats are able to maintain consistent levels of coordination and coordinated switching upon reversal of outcome contingencies. Dyads appear to rapidly adapt their behaviour losing a preference for mutual cooperation over mutual defection in the session directly following the reversal. Over sessions

however they alter their preference towards the re-assigned cooperation arm and show an increase in mutual cooperation (R) trials and a decrease in mutual defection trials (P). The change in the propensity to engage in mutual cooperation and mutual defection indicates that dyads remain sensitive to outcomes after multiple sessions of high mutual cooperation in the standard condition and are able to re-establish similar high levels of cooperation and mutual cooperation following a reversal.

3.4.1 Isolated Condition

In the standard iSH condition, each animal had access to ongoing social information as they make choices as well as information from the history of received rewards. A central question here is the extent to which cooperative and coordinated behaviours depend upon this social information. To examine the role of proximate social information in supporting these behaviours, we conducted the following isolated version of the standard iSH task.

The isolated condition is identical in every way to the standard condition except that in this case the two mazes that were previously adjacent are now entirely separated removing all proximate social cues. The structure of each T-maze and the experimental room is such that each maze is visually occluded from the other and it is extremely unlikely that either animal could distinguish the choice of its opponent by any auditory cues that may emerge from the operation of either maze. Unlike the reversal condition, the designated cooperation and defection arms remain unchanged. It is important to note that by isolating each animal within a dyad so that their choices are made privately is closer to how game theory

is intended to be applied, as now each choice is made independently and without the presence of social information.

3.4.2 Isolated Condition Results

Six dyads from the standard condition, after completing fourteen sessions were subsequently exposed to this new isolated condition for a further twelve sessions. Here we saw a moderate but significant drop in simple coordination directly following the transition to the isolated condition (paired t test $t_{14}=3.2643$ $p<0.05$) (fig3.4.1(a)) though this still remained greater than chance (one tailed t-test $t_5=2.0826$ $p<0.05$). Following this initial decrease in coordination, a simple regression analysis revealed a moderate increase over sessions ($r^2_{0.5531}$, $f_{13.4062}$, $p<0.05$) that is overall greater than chance ($t_{71}=20.982$ $p<0.001$) (fig(3.4.1c)). When we now examine the proportion of coordinated switching trials, we see that in the absence of social information, the likelihood of coordinated trials occurring when one member of the dyad changes their choice became indistinguishable from chance (fig3.4.2b and c). However this difference was only clearly observable from the second rather than first session ($t_{14}=2.5046$ $p<0.05$).

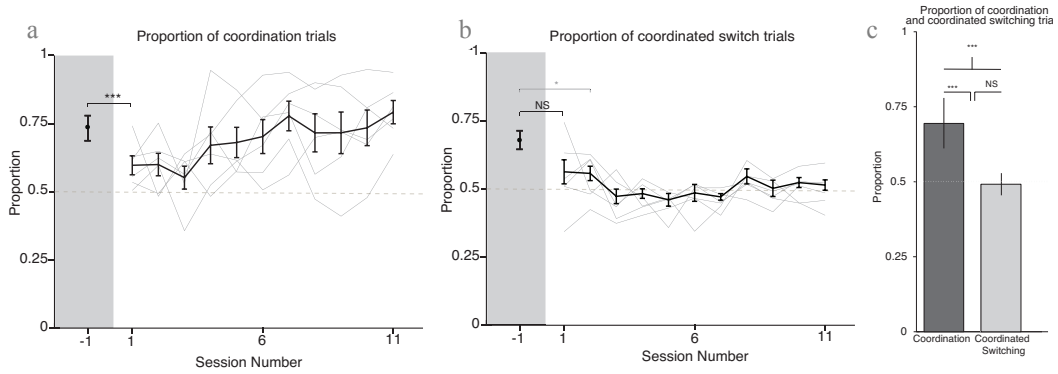


Figure 3.4.1 | Dyads maintain simple coordination & decrease coordinated switching in isolation

Proportion of total trials where each agent makes the same choice across session (a) across all sessions, conditional probability of coordinating given either member of a dyad changed choice from the prior trial (b) and mean proportions of coordination, and coordinated switching across all pairs and switching (c). Light grey lines indicate pairs black line the mean, and error bars are SEM.

This high level of coordination but not coordinated switching suggests that active coordination of trial-by-trial choices required ongoing social information, rather than a model of the other's choices. Simple coordination here was likely maintained by the development of shared preference.

In both, the all equal and standard iSH conditions we observed a robust tendency to engage in alternation. This may have been a response to the T-maze design of the assay, or upon social information. Isolated animals also show a similar tendency to alternate as in the prior condition, however they show a greater likelihood to perseverate in defection than in cooperation over sessions, indicating that the proximity of a conspecific is not sufficient to produce alternation (fig 3.4.2). Rather, it suggests that under these payoff conditions, the presence or absence of conspecific plays an important role on whether perseveration is increasingly directed towards cooperation or defection, respectively.

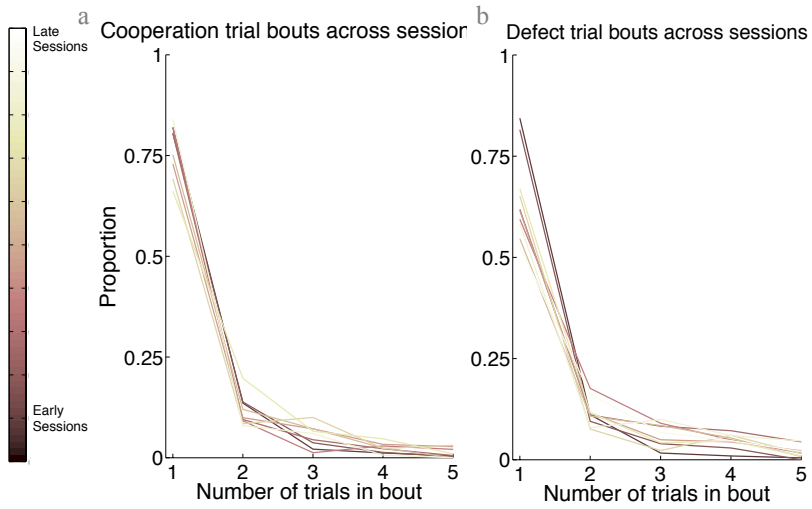


Figure 3.4.2 | Cooperation and defection bout length across conditions. Probability density functions for the number of continuous cooperation (a) and defection (b) trials across all rats, Color coded by session, darker to light for early to late sessions.

Examining the proportions of mutual cooperation and defection trials, we see from the first session following the transition from the standard to the isolated condition, there is a noticeable decrease in the proportion of mutual cooperation trials in the following sessions (paired sample t_{14}=3.329 *p<0.05*) (fig 3.4.2.). A similar, though non-significant increase is also present in mutual defection trials (paired sample t_{14}=-0.2758 *p=0.0505*). Over sessions, we see a moderate increase in mutual defection trials ($r^2_{0.5723}$, $f_{29.9695}$, *p<0.01*) and a moderate decrease in mutual cooperation trials ($r_{20.1927}$, $f_{11.3151}$, *p<0.01*).

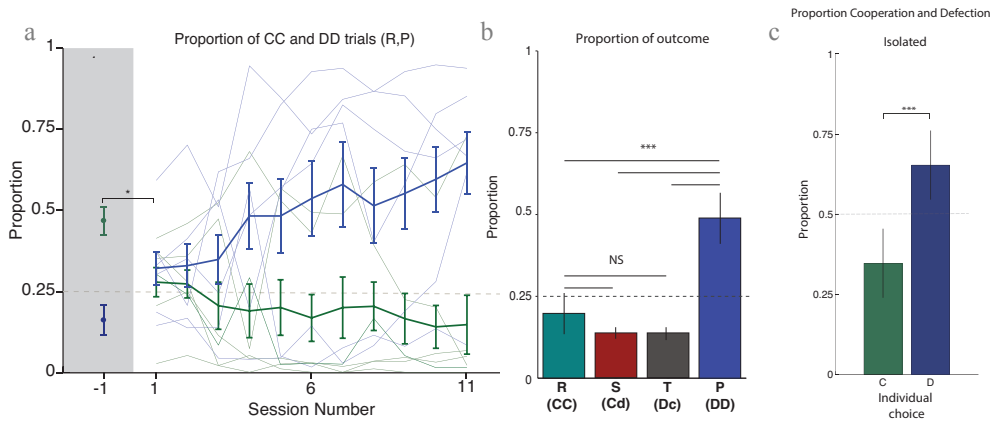


Figure 3.4.3 | Dyads increasing engage in coordinated cooperation. Coordinated cooperation; Reward trials (green) and coordinated defection, punishment trials (blue) across session (a) and all outcomes across sessions and pairs, reward (green), punishment (blue), anti-coordinated trials, temptation (black) and sucker (red) (b). (c) mean proportion of cooperation and defection trials across all rats and sessions. Error bars are SEM.

Across all sessions and dyads (fig 3.4.3(b)), we see essentially the opposite pattern for R and P outcomes than that seen in the standard condition. Here, only P outcomes were greater than chance (one sample t test t_{65}=6.2470 *p<0.001*). A repeated-measures ANOVA revealed a significant effect of outcome upon proportion of choice ($F_{3,43.42}$=1.705 *p<0.0001*), post hoc analysis by Tukey Cramer HSD indicated that the mean proportion of P outcomes (M=0.4895, ±0.1021) was greater than that of R outcomes (M=0.1966 ±0.0244) and anti-coordinated outcomes (M=0.1770 ±0.0941). This change towards increased defection was also represented in the proportion of all rats' cooperation and defection choices. We now see that the proportion of individual defection trials was greater than chance (t_{143}=7.0015 *p<0.001*) and, necessarily, than cooperation trials.

It is important to note that the observed increase in mutual defection trials and decrease in mutual cooperation is true of the population as a whole. However, two of the six dyads behaved quite

differently. One maintained a high level of mutual cooperation throughout all isolated sessions and a further dyad maintained high levels of mutual cooperation until the final two sessions where they then increased in mutual defection. Interestingly, neither of these pairs showed a correspondingly high level of coordinated switching which may indicate that both pairs simply maintained a strong preference for cooperation that was established in the prior sessions in the standard condition.

Thus far, we have seen that in the absence of social information, the majority of dyads showed high rates of defection leading to increased mutual defection; the risk dominant Nash equilibrium. This was matched by a rapid decrease in mutual cooperation trials. This change led to a direct decrease in received outcomes in the session following transition to the isolated condition (paired sample t_{14}=-2.4875 $p<0.05$) which does not change over sessions. This increase in risk dominant over payoff dominant choices by dyads maintained a level of reward greater than would be achieved from each rat choosing randomly (one tailed t test t_{71}=5.3557 $p<0.0001$), yet less than would have been achieved from following a complete all defect strategy (one tailed t test t_{71}=-6.4073 $p<0.001$) or that achieved in the standard condition (paired t test t_{210}=-8.2476 $p<0.0001$).

Notably, despite changing to risk dominant rather than payoff dominant strategy, the proportion of maximum rewards received over sessions does not increase (fig 3.4.4). This means that dyads demonstrated a change in strategy from high levels of mutual cooperation previously to mutual defection in the isolated condition, yet this change did not lead to a change in received outcomes. Changing from the payoff dominant to risk dominant choices without a change in received outcomes may indicate that

rats in this experiment are either in general, or in isolation, risk averse and that social information plays an important role in overcoming this risk aversion, allowing an increased frequency of risky cooperate choices.

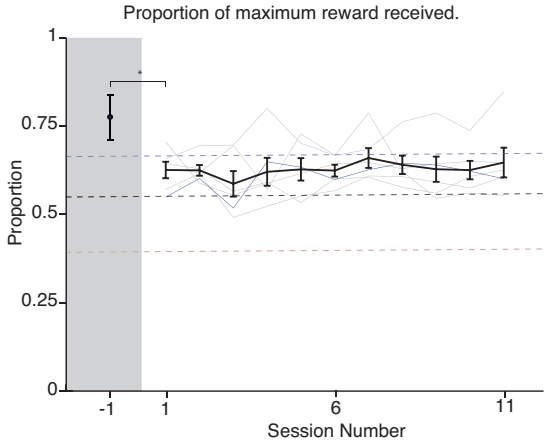


Figure 3.4.4 | Dyads achieve reward consistent with risk dominant payoff.
Proportion of outcomes, the number of pellets received by the dyad as a proportion of the maximum reward available from an all cooperate strategy. Dashed lines indicate the proportion of reward yielded from perfect 'all defect' (blue), random (black) or anti coordination (red) strategies. Black line is mean, brick lines are individual dyads and error bars SEM

3.4.3 Summary

To examine the role that ongoing social information has upon the behaviour of dyads engaging in the iSH task, we transitioned six dyads to an isolated condition where all social information was removed. The above data indicate that social information plays a key role in three fundamental aspects of dyads' behaviour.

Firstly, while simple coordination decreases directly following the transition from the standard condition, it increases across sessions towards levels similar to that observed in the standard condition. In contrast,

coordinated switching rapidly decreases until it is indistinguishable from chance. This is particularly important as rats maintained a similar preference for alternation as they did in the standard iSH condition. This demonstrates that the capacity to synchronize trial-by-trial choices is dependent upon ongoing social information while the tendency to alternate is not. Simple coordination, choosing the same side, quickly recovered to similar levels as a result of an increasing and shared preference for mutual defection. This indicates that while animals could coordinate, they were doing so only on average, rather than in a trial-to-trial manner, increasing the likelihood of receiving S outcomes for cooperation.

It is important to note that in the standard condition, social information was available to both members of a dyad both, preceding either rat's choice and after choices were made. In the Isolated condition, this information is absent in both cases. Each rat can, in principle, determine the choice of their opponent following a cooperation choice by whether they receive the R or S outcome, an R outcome if their opponent cooperated and an S outcome if they did not. The two outcomes available for defection, T and P, are identical and therefore provide no information as to whether the partner of a rat that has defected had also done so. This means that the rat in this condition cannot distinguish between the importance of social information preceding choice and following it.

Secondly, the previously evidenced preference of dyads to engage most frequently in mutual cooperation, consistent with the payoff dominant equilibrium, in the standard condition is replaced in this condition with a tendency to engage in mutual defection, the risk dominant equilibrium. This suggests that ongoing social information plays a central

role in maintaining mutual cooperation. Another interpretation of this data would be animals chose most frequently the option where outcome was not dependent upon the other's choice. However two of six dyads were able to maintain high levels of mutual cooperation showing that social information is not necessarily required for high levels of mutual cooperation.

Here the quantity of pellets received on average was lower, as would be expected by a decrease in R outcomes, than in the standard condition. However the change in strategy to one that increases mutual defection over mutual cooperation did not increase the amount of rewards received. This indicates that if dyads had maintained the same tendency to cooperate throughout the isolated condition as they did in the early sessions they would have not decreased the total amount of received reward in a session. This suggests that rats in this task may be risk averse, and prefer a constant moderate reward yielded by T or P rather than the possibility of receiving an S outcome on some trials, even if the overall proportion of reward received in a session remains the same.

The low rates of coordinated switching demonstrated in this condition reduce the likelihood of each rat cooperating on the same trial as their opponent. This makes choosing to cooperate a more risky proposition. Also the presence of conspecific may reduce this risk aversion via social buffering or social facilitation. The data here does not distinguish between less favorable outcome statistics for cooperative choices, or more social influence upon risk perception mediated by the proximity of a conspecific.

3.5.1 Decoupled Condition

The removal of social information from previously cooperating dyads led to a rapid increase in both unilateral and mutual defection. It remains unclear, however, in what manner social information is being used to support cooperation. Data from the standard condition indicates that presence of social information supports trial-by-trial coordination of behaviour. This leads to outcome statistics that are more favorable towards cooperation in that the likelihood of receiving an R outcome on a particular trial was higher. This goes some way to mitigating the risk of cooperation. However it remains unclear if this difference in preference for more risky cooperate choice is supported by the presence of conspecific, the average likelihood of receiving an R rather than S outcome for cooperation, or whether it depends upon the content of social interactions between animals.

To determine whether the presence of a conspecific and the average likelihood of receiving an R rather than S outcome is sufficient to support cooperate, we implemented the following Decoupled reward condition. In this condition, dyads were provided the probability of receiving an R outcome on the cooperate arm as animals did in the late sessions of the standard iSH condition except here, the outcome each animal receives is no longer contingent upon the choices of their opponent. In practice, this means that an outcome of 3 pellets is available on the cooperate arm for both mutual cooperation (R) and unilateral cooperation (S) with a likelihood of 0.8. In the remaining cooperative trials an outcome of 0 was provided. The outcome received for either unilateral (T) or mutual defection (P) remains a constant moderate reward of 2 pellets as before.

	Cooperate	***Defect***
Cooperate	3/0*,***3/0*** (**R**,***R***)	0/3*,***2 (*** ***S*** ,***T)***
Defect	2,***0/3*** (**T**, ***S)***	2,***2 (***P,***P)***

*80% of trials 3 / 20% of trials 0

3.5.2 Decoupled Condition Results

Here 5 naive dyads were exposed to the conditions detailed above for 8 sessions. As in prior conditions, we first examined the proportions of coordination and coordinated switch trials. It is important here that each animal coordinating its choice with its opponent has no influence upon received outcome. Across all sessions, we saw that the proportion of coordination and coordinated switch trials occur greater than chance (t_{39}=9.5790 *p<0.001*, t_{39}=8.8365 *p<0.001*) but, similarly to the all equal condition they are indistinguishable from each other (fig 3.5.1a). As in the standard and all equal condition, we see no significant changes over sessions.

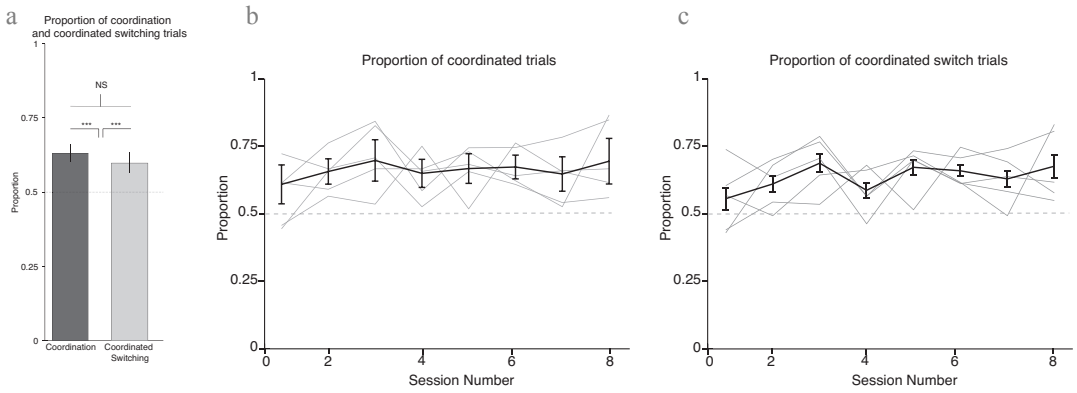


Figure 3.5.1 | Dyads continue to coordinate when outcomes are not contingent upon collective choices.
(a) Mean proportions of coordination, and coordinated switching across all pairs and sessions. Proportion of total trials were each agent makes the same choice by sessions (b) and the conditional probability of coordinating given either member of a dyad changed choice from the prior trial (c). Light grey lines indicate pairs black line the mean, and error bars are SEM.

This data indicates that in the presence of different outcomes, even when they are not contingent upon collective outcomes, rats still demonstrate a consistent preference for coordinated choices even when this does not increase their received outcomes.

Our central question here is the extent to which animals cooperate, and the extent to which dyads engage in mutual cooperation. Across all sessions and dyads, a repeated measures ANOVA revealed a significant effect of outcome upon proportion of choice ($F_{3,42.78}$=0.3297 *p<0.0001*), post hoc analysis by Tukey Cramer HSD indicated that the mean proportion of P outcomes (M=0.3280, ±0.013) and R outcomes (M=0.3293 ±0.0126) were significantly greater than anti-coordinated outcomes (M=0.1714 ±0.0134) but not different from each other (fig 3.5.2a). Over

sessions we see no significant change in either mutual cooperation or mutual defection trials. Across all sessions we see that animals are equally likely to cooperate as to defect (fig 3.5.2c).

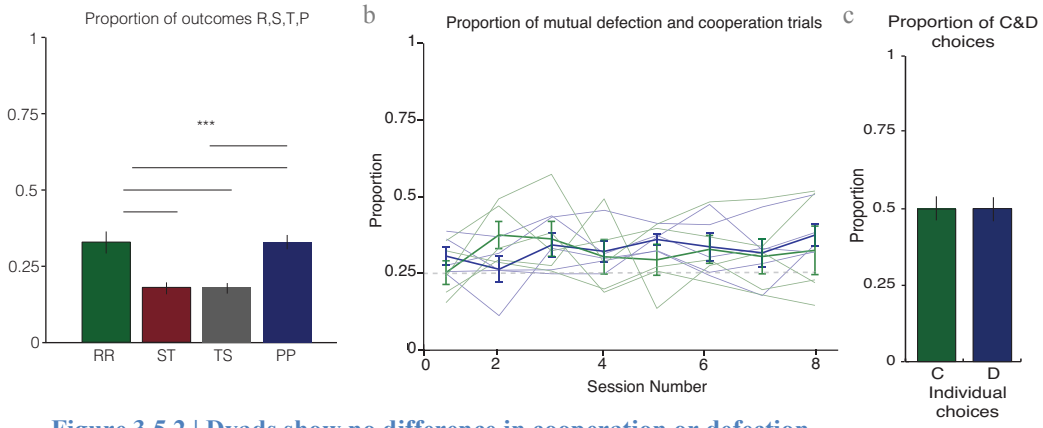


Figure 3.5.2 | Dyads show no difference in cooperation or defection.
Mean proportion of mutual cooperation (green) mutual defection (blue) and two anti coordination type (red and black) (a). Progression of mutual cooperation (green) and mutual defection (blue) across sessions (b) and proportion of cooperation (green) and defection (blue) trials over sessions. Error bars are SEM. Grey dashed lines indicate chance.

3.5.3 Summary

This decoupled condition provided animals with the same outcome statistics for cooperation and defection encountered by animals in later sessions of the standard condition where mutual cooperation was highest. However in this case we removed the contingency between *collective* choices and outcome. Despite there being no benefit to coordination, animals demonstrated a tendency and a capacity to coordinate robustly across sessions. This supports findings from the equal reward condition where animals have a preference for coordinating their choices regardless of outcome.

The central question posed by this experiment was whether the rates of cooperation seen in the standard condition were dependent purely upon the high likelihood of receiving a high outcome on the cooperate arm. Here, even though that likelihood is the same, we see that dyads do not develop a preference for mutual cooperation nor individual animals, on average, for cooperation. The iSH task is designed to emphasize that cooperation requires that each animal knows that the choice of the other determines their outcome on the cooperate arm and that their opponent is likely to cooperate. Here we have removed both those requirements and simply ask if a high likelihood (0.8) of receiving the maximum outcome on the cooperate arm is enough to promote cooperation. The crucial difference is that even though the likelihood of cooperation yielding a high outcome is the same, each animal has no information on whether on a particular trial they will receive the high reward or not. That animals show no preference for cooperation under these conditions indicates that the reward statistics observed in the standard condition are also insufficient to produce high levels of cooperation. Further, the presence of a conspecific in itself is also not sufficient to mitigate the risk of the cooperate arm. This suggests that it is not simply how likely a high versus low outcome will be received for cooperation on average, but rather upon the ability to leverage ongoing information in the interaction with a conspecific prior to choosing to generate an expectation that a cooperative choice will be rewarded.

3.6.1 Cross condition comparisons

In the each of the experiments that we exposed animals to the iSH payoff matrix, the standard and isolated conditions, there appeared to be a learning period and an asymptotic period where animals expressed stable preferences. However under the modified payoff conditions there were no

notable changes in behaviour over sessions. We will now proceed to compare the later 8 sessions in the isolated and standard iSH condition where behaviour appeared to have stabilized and all sessions from the decoupled and equal reward condition.

As animals tended to alternate in each condition under the standard iSH payoff conditions, coordinating this alternation measured by coordinated switching determines how often cooperation trials lead to an R rather than an S outcome. Social information is required for coordinated switching. A one way ANOVA across conditions revealed a significant effect of condition upon coordinated switching ($F_{3,45.75}$=0.34637 *p<0.0001*), post hoc analysis by Tukey Cramer HSD indicated that coordinated switching was greatest in the standard iSH (M=0.6609 ±0.097) and decoupled condition (M=0.6573 ±0.0138) which were both significantly greater than the equal rewards condition (M=0.5576 ±0.0126). Coordinated switching was greater in all conditions with social information than the isolated iSH condition (M=4.954 ±0.0126). Only in the isolated iSH did coordinated switching not exceed chance ((one sample T test) standard: t_{79}=141.580 *p<0.001,* equal rewards: t_{47}=5,1858 *p<0.001*, decoupled: t_{39}=8.8365 *p<0.001*). This indicates that animals require social information to coordinate trial-by-trial alternation in choices, but not economic incentives, as animals do so greater than chance in the equal rewards condition and in decoupled condition.

A one-way ANOVA upon the proportion of mutual cooperation (R) trials in the asymptotic period confirmed a significant effect of condition (($F_{3,54.7908}$=1.4438 *p<0.0001*) (fig 3.5.1b). Mutual cooperation was highest in the standard iSH condition (M=0.5458, ±0.0181) and lowest in the

isolated iSH condition (M=0.1760, ±0.0234) while the equal rewards (M=0.2960, ±0.0282) and decoupled conditions (M=0.32860 ±0.02587) were equivalent. Further emphasizing the preference for the risk dominant equilibrium, mutual cooperation in the isolated iSH condition was lower than chance (t_{79}=-3.3063, *p<0.001*).

Mutual defection (P) in the asymptotic period of each condition followed a similar pattern, also showing a significant effect of condition (($F_{3,38.05}$=1.2208 *p<0.0001*), (fig 3.5.1c). Mutual defection was highest in the isolated condition (M=0.5483 ±0.0259) and was greater than all other conditions (all equal: M=0.2733 ±0.0259, standard iSH condition: M=0.2054 ±0.0200, and decoupled condition: M=0.3293 ±0.0283). Animals in the standard condition were also less likely than chance to engage in the non preferred Nash equilibria, mutually defecting less than chance (one tailed t test t_{79}=-3.3063).

Coordinated switching was greatest in the standard iSH and decoupled condition. This supports the importance of the role that this trial-by-trial coordination has in supporting coordination, as in the only case where it did not exceed chance, animals reliably defected. However it is not sufficient to explain high rates of cooperation as in the decoupled condition animals demonstrated the same proportion of coordinated switching as in the standard condition, yet also did not have any economic incentive to do so. That coordinated switching is higher in the decoupled condition may be a response to outcomes being unpredictable, still when

rewards were not contingent upon collective action.

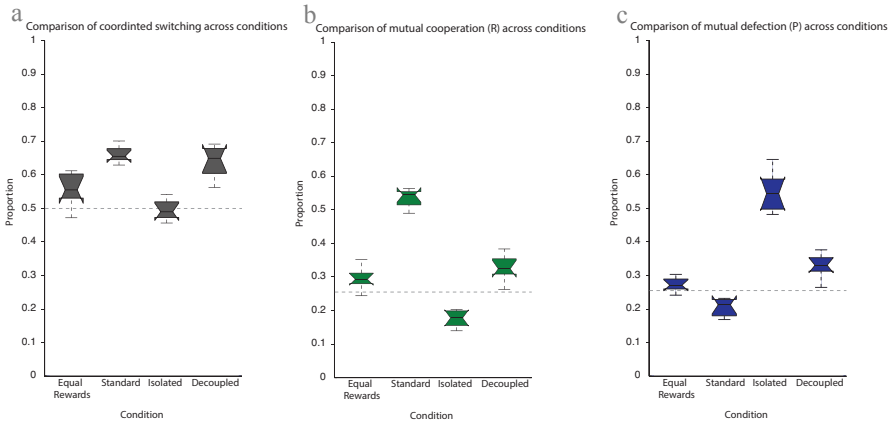


Figure 3.5.1 | Comparison of coordinated switching and mutual cooperation and defection across all conditions. Notched box plots indicated the mean proportions of coordinated trials (a) mutual cooperation (b) and mutual defection (c) from the asymptotic period of each condition.

3.6.2 Summary

The iSH has two Nash equilibria, one payoff dominant, where each player risks forgoing reward, for a large reward, and one risk dominant, where each player chooses a safe moderate reward. These equilibria, mutual cooperation and mutual defection are collective choices from which neither player can increase their outcome unilaterally by choosing differently. Animals in all conditions display a consistently high level of alternation showing that they do not tend to perseverate in one or the other equilibrium as the concept predicts. Nonetheless, in both the standard and isolated condition we do see a preference on average for the payoff and risk dominant equilibria respectively. In the two cases where rewards were determined by the iSH payoff matrix, animals showed a strong preference for behaviour consistent with their preferred equilibrium, mutual cooperation in the Standard iSH condition and mutual defection in the isolated iSH condition. Between these two conditions this preference, at

least in part, seems determined by their capacity to engage in coordinated switching. Animals coordinate this alternation no greater than chance in absence of social information, decreasing the likelihood of achieving the payoff dominant equilibria, and thusly reducing the value of cooperation. In the standard iSH, high levels of coordinated switching make achieving the payoff dominant equilibrium more likely, which supported cooperation. However, animals in the decoupled condition showed the same propensity to engage in coordinated switching. In the decoupled condition the R outcome provided 2.4 pellets on average, making cooperation more rewarding on average than defection. Therefore, to maximize reward each animal should increase cooperation. Maintaining coordinated alternation with a partner that does not develop this preference for cooperation and is therefore costly. One explanation for this would be that the benefit of cooperation is not sufficient to drive a preference for it, however, they do develop this preference when outcomes are contingent upon collective action. This suggests that animals may be motivated to coordinate their alternation when rewards are unpredictable.

3.6.3 Analysis of switching.

The ability of animals to coordinate in a context of alternation appears to be a fundamental factor in determining dyads' preference for mutual cooperation or mutual defection and is mediated by the presence of social information. Coordinated switching is highest when animals receive variable rewards on the cooperate arm suggesting that prior received outcomes are important in determining this behaviour. We therefore examined the conditional probability of switching choice given the

outcome of the prior trial from the asymptotic period of each condition and the changes in these conditional probabilities over sessions.

In the conditions where rewards are not contingent on collective action, the equal rewards and decoupled condition, coordination is most likely to be followed by a change in choice. First, examining the rate of switching after any outcome using a one way ANOVA showed a significant effect of condition ($F_{e,11.8116}$=0.852, p<0.0001). Post hoc analysis by Tukey Cramer HSD revealed that animals in the equal rewards condition had the highest rate of alternation (M=0.8110 ±0.0094) again indicating that high levels of alternation are not elicited in response to reward difference, but rather interacting animals may decrease their alternation as they navigate the different available outcomes. A one-way ANOVA revealed a significant effect of outcome upon likelihood of switching ($F_{3,58.81}$=0.852, p<0.0001). Post hoc analysis by Tukey Cramer HSD showed that in the absence of reward differences, animals' likelihood of switching was greatest following coordinated choices (R:M=0.9205, ±0.072, P:0.9476, ±0.0188) and least following anti-coordinated outcomes (S:M=0.6758,±0.097 T:M=0.6940, ±0.0832) demonstrating that the proximity of a conspecific has a strong influence upon each animal's likelihood of switching. In the decoupled condition where rewards are probabilistic but favor cooperation, we again see a significant effect of prior trial upon the likelihood of switching ($F_{3,65.32}$=0.906, p<0.0001), where animals are most likely to switch following a coordinated (R:M=0.7907, ±0.0251, P:M=0.9101, ±0.0124) rather than an anti coordinated trial (S:M=0.5647 ±0.045, T:M=0.6564 ±0.0631). This is further modulated by whether the prior trial was cooperative. Animals were less likely to switch following a cooperative coordinated trial (R)

than a defective coordinated trial (P) and less likely to switch following unilateral cooperation (S) than unilateral defection (T).

In the two conditions where rewards are contingent upon collective choice, animals now become less likely to switch their choice following their established preference for cooperation and defection. In the standard iSH condition ($F_{3,89.986}$=4.352, p<0.0001), animals are less likely to switch following cooperation (R:M=0.647 ±0.0536, S:M=0.451 ±0.0962) than defection (P:M=0.952 ±0.065, T=0.825 ±0.086). In each case they are less likely to switch following a unilateral than a coordinated choice, indicating that while reward is now the primary predictor of switching, it is also modulated by whether their opponent made the same choice. Unexpectedly, rats in this condition are least likely to switch following an S outcome. Across conditions this is the only case where animals are less likely than chance to switch (t_{79}=-3.3442 p<0.001). In the isolated context ($F_{3,85.986}$=4.352, p<0.0001) however, animals are least likely to switch following defection, regardless of whether it was coordinated (P:M=0.4699 ±0.125) or unilateral (T:M=0.4830 ±0.0239). In isolation, animals cannot distinguish between unilateral and coordinated defection as the rewards are the same. However, they are also equally likely to switch following an R (M=0.8541 ±0.072) or an S (M=8039 ±0.062) outcome. Only in the isolated condition, therefore, does the likelihood of switching not depend upon the behaviour of their opponent.

The only case where animals are less likely than chance to change their choice is after receiving an S outcome was in the standard iSH condition. While each condition has its own distinct pattern of switching given prior outcomes, this is perhaps the most striking finding as it

indicates a tolerance to being suckered and thus deviates, at least at first glance, from economic expectations. Such tolerance, however, would suggest a mechanism by which animals overcome their tendency to alternate and establish consistent mutual cooperation. To determine whether this tendency to tolerate S outcomes changed over sessions, we conducted a simple regression analysis upon the conditional probability of switching and S outcomes in the prior trial in each condition. Only the standard iSH condition showed a significant change over sessions, with animals being increasingly less likely to switch following S trials ($r^2_{-0.8724}$, $f_{15.9577}$, $p<0.0001$). A similar decrease in the likelihood of switching following an R trial was also present ($r^2_{-2.1302}$, $f_{19,3702}$, $p<0.0001$) indicating that animals are increasingly likely to perseverate following a cooperation trial. In the isolated condition there was a similar decrease, over sessions, in the likelihood of switching following each defection outcome, (T: $r^2_{-0.2960}$, $f_{10.3788}$, $p<0.0001$), (P: $r^2_{-0.4553}$, $f_{8.7328}$ $p<0.0001$). However, as we saw in the asymptotic sessions, there is no significant difference between their tendency to switch following T or P trials when there was such a difference between S and R outcome in the standard condition. This indicates that in the standard iSH case, the manner in which they biased their alternation was dependent upon the choice of their opponent while in the isolated iSH condition, it was not.

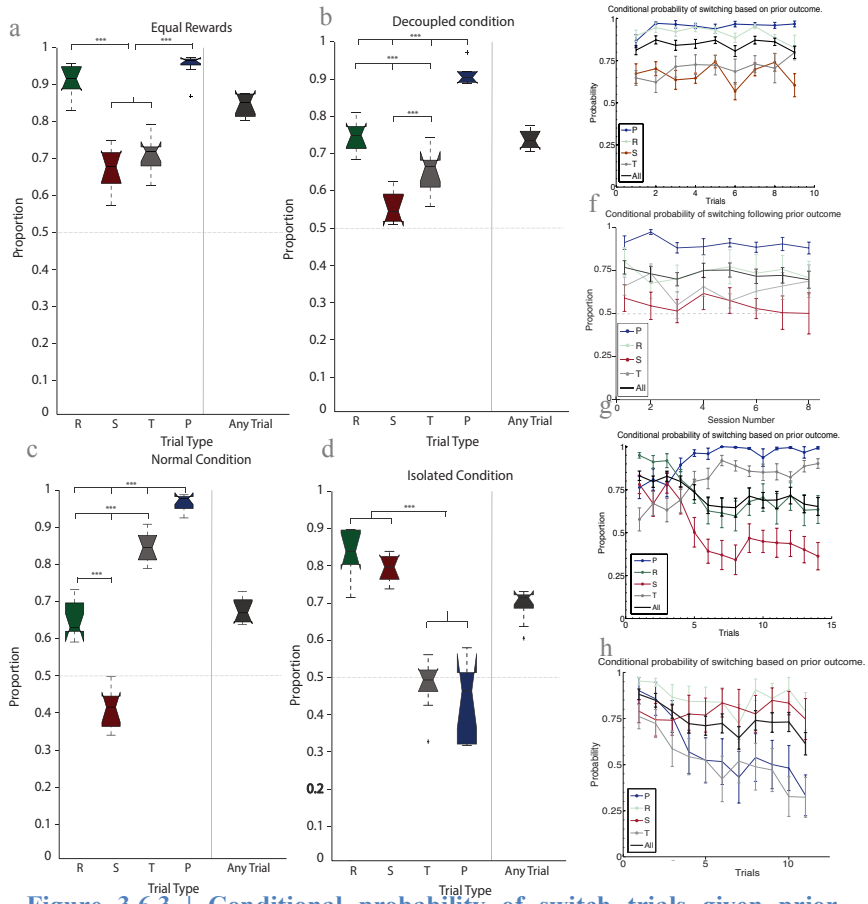


Figure 3.6.3 | Conditional probability of switch trials given prior outcome.

Notched box plots showing the mean conditional probability of 'switching' given outcome on prior trial for equal rewards (a,e), decoupled (b,f), standard (c,g) and isolated conditions (d,h) from the asymptotic period. (e-h) show this same data across all sessions for each condition (R=green, Red=S Grey=T, Blue=P and black indicates all). Error bars are SEM.

3.6.4 Summary

To determine precisely what was influencing this switching behavior, we compared the likelihood of a switch trial following events in the prior trial. In the standard condition and decoupled conditions where social information and outcomes are uncertain, both social information and

reward influence the likelihood of switching. In the decoupled condition where rewards are probabilistic, the animals' likelihood of switching was primarily influenced by whether the prior trial was coordinated, then by whether it was either a cooperate or defect trial. When rewards were contingent upon collective action, switching behaviour was most influenced by whether they cooperated or defected in the prior trial, and then by whether the trial was coordinated or otherwise.

Removing social information resulted in the likelihood of switching being determined by each animal's prior choice rather than the behaviour of the other animals or resulting outcome in the previous trial. That outcomes do not seem to influence behaviour, is a surprising finding though this may be accounted for by being averse to the risky outcomes on the cooperate arm and therefore a preference to defect regardless of whether a prior cooperation choice was rewarded or not.

Alternation was highest in the all equal condition, and the same in each of the other conditions. This indicates that animals reduce their tendency to alternate when cooperation and defection provide different or uncertain rewards. This may provide an important constraint upon animals' capacity to coordinate their preference for cooperation or for defection in the standard and isolated iSH conditions.

Animals in the standard iSH condition display a unique tendency to tolerate low S outcomes compared to the other conditions. It is also the only condition in which the likelihood of switching following an S outcome decreases over sessions. Animals in this condition were extremely likely to switch following the corresponding T trial, which

suggests that perseveration after an S outcome is likely to lead to the highest reward (R) in the subsequent trial. This suggests that tolerance of S outcomes is likely to support both cooperation and coordination.

3.6.5 General Linear Models and choice history

Events in prior trials modulate the pattern of alternation of rats in each condition. However, high levels of alternation mean that examining only the influence of the trial directly preceding choice is necessarily limited. To address how the history of events each animal has experienced influences future choices, we generated general linear models to examine the role of cooperative outcomes and the role of defection outcomes on the propensity to cooperate. We then created a more complete model that attempts to examine the influence key components of choice. In each of these models, we ask how the prior choices and outcomes from the last seven trials affect the rat's choice on the current trial.

To do this, we imported the data into R (R Core Team, 2016) and used both standard Logistic (Generalized) Linear Models and L1 penalized Logistic Regression using the elastic net approach. In brief, we hoped to use the L1 penalty from the elastic net GLM to allow larger models with more predictors from past trials while not over-fitting or doing extensive model comparison. The L1 norm effectively places a prior that many coefficients in the model are likely to be zero, which should be true for the predictive value of events that occurred many trials in the past.

We modeled rats as individuals choosing to cooperate or defect. The rat's choice on each trial was predicted using logistic regression in

three separate models. The predictor design matrix was constructed by taking the value of each predictor from one to seven trials into the past (column one is the predictor one trial back, column two is the same predictor two trials back, and so on) and combining them to create the full model design matrix. No constant term was added. We included session boundaries so that it was not possible for lagged predictors to cross from one session to another.

Here, for historical reasons and to ensure a learning period with sufficient sessions in each case, we included the first five sessions from each condition, including decoupled and isolated conditions. The asymptotic period included the remaining sessions in each condition. To more closely examine the effects of the history of cooperate trials, we generated a model (eq 3.1) that consisted of prior cooperation choices split into R and S outcomes.

$$\ln\left(\frac{P(Cooperate_t)}{1-P(Cooperate_t)}\right)=\beta_1 Reward_{t-1}+\ldots+\beta_7 Reward_{t-7}+\beta_8 Sucker_{t-1}+\ldots+\beta_{15} Sucker_{t-7} \quad (3.1)$$

We examined each of these outcomes seven trials into the past for each of the four conditions. The model was run using the standard GLM in R and using the glmnet package with cross-validation to select the penalty weighting, which maximized out-of-sample prediction. We selected the lambda (weighting parameter) as one standard deviation out from the cross validated minimum error, as suggested by the package authors to avoid over penalization. Positive values for each predictor indicate that they positively predict cooperation, negative values predict defection.

Across all conditions the tendency to alternate is clearly visible, with both R and S outcomes negatively predicting cooperation in the following trial. It is important to note that in only the standard and isolated iSH conditions are the R and S outcomes clearly defined and distinct. In the decoupled condition it is possible that an animal was rewarded on an S trial or not on an R trial and vice versa. In the equal rewards condition the received rewards are, of course, identical (fig 3.6.4).

Comparing the learning and asymptotic period in the standard condition we see that, during the learning period, receiving R outcomes increased the likelihood of cooperation in later trials and this effect is more pronounced in the later sessions. The clearest change between learning and asymptotic periods can be seen in the influence of S outcomes. Previously we saw that in this condition animals are most likely to perseverate in cooperation following a low S outcome. In the learning period we see that that sucker trials marginally predict future cooperation. However, in the asymptotic period S outcomes positively predict cooperation up to five trials into the past. This again supports the suggestion that in important aspect of animals increasing mutual cooperation is in part due to a learned tolerance to S outcomes.

In the isolated condition, we can see a clear effect of received outcomes upon cooperation. In the learning phase, when animals have recently transitioned from the standard iSH condition, we see that high R outcomes positively predict cooperation. R outcomes continue to predict future cooperation in the asymptotic period. S outcomes on the other hand have a marginal influence upon cooperation following transition to the isolated condition suggesting a reduction to tolerance to S outcomes in the

absence of a conspecific. However in the asymptotic period, S outcomes now decrease the likelihood of future cooperation. Animals in the isolated react to outcomes in a more conventional manner, with positive outcomes increasing the likelihood of cooperation and negative outcomes decreasing it. In isolation, previously tolerated S outcomes instead predict future defection. This may result from animals failing to coordinate their alternation in this condition, meaning that they received S outcomes more frequently that R outcomes when cooperating. It may also be due to the absence of the conspecific.

We next examined the effects of the history of defect trials upon cooperation split into T and P outcomes applying the same conditions as in the prior model (eq.3.2).

$$\ln\left(\frac{P(Cooperate_t)}{1-P(Cooperate_t)}\right)=\beta_1 Punishment_{t-1}+\ldots+\beta_7 Punishment_{t-7}+\beta_8 Temptation_{t-1}+\ldots+\beta_{15} Temptation_{t-7} \quad (3.2)$$

Here, what is being compared is the role of the history of coordinated and uncoordinated defection on subsequent choice (fig 3.6.5). Unlike with cooperation outcomes, defection in all conditions yields the same outcome regardless of whether it was coordinated or not. As this might predict, we see a much more similar influence of both T and P outcomes across all conditions in the learning period. In the isolated condition where animals had a preference for mutual defection, prior P outcomes positively predicted defection. In the standard iSH condition where they had a preference for mutual cooperation, this effect was much reduced. Unilateral defection (T) however had little effect upon future cooperation in either condition.

In the prior two models, we split the cooperation and defection outcomes into separate models as including possible trial types leads to co-linearity in the model. While this allows us to explore the effect of both cooperation outcomes and both defect outcomes, we cannot compare across models. Also while the effect of alternation is clear in the predictive value of each outcome, it is also very likely to influence trials further into the past. To address this, we generated a third more comprehensive model and organized the parameters in the following manner (eq.3.3):

$$\ln\left(\frac{P(Cooperate_t)}{1-P(Cooperate_t)}\right)=\beta_0 CoordinateSecond+\beta_1 Choice_{t-1}+Choice_{t-1}+\ldots+\beta_5 Choice_{t-5}+\beta_6 Pellets_{t-1}+\ldots+\beta_{11} Pellets_{t-5} \quad (3.3)$$

Here, we examined the effect of going second (coordination with knowledge), prior choices, T outcomes and the relative gain or loss in pellets by taking the risky cooperate choice. This model used the other animal's choice, coded as 1 for cooperate and -1 for defect when the modeled animal made his choice second and 0 otherwise. The prior choices were coded as -1 for defection and 1 for cooperation, indicating stay preference when the predictors' coefficients are positive and a switch tendency when they are negative. The prior rewards were coded relative to the safe option, defect, which always produced 2 pellets. Cooperate choices were therefore coded -2 when they received no pellets and +1 when they received three. Choices to the defect arm were coded as 0, two minus two, the reference amount.

As this model places greater requirements upon the data, we will focus on the standard and isolated condition, which each consisted of more sessions and more trials. However, as before, the findings from the other two conditions are also shown.

Here, we see that in the standard condition there were several differences between the learning period and asymptotic period. In this condition, rats show an increased influence of second choice upon coordination (fig 3.6.6). Examining the tendency to alternate, we see that in the asymptotic period, while there is an increased tendency to change one's choice in the trial directly following, animals show consistency in choices more trials into the past. Having captured much of the tendency to alternate in this parameter, we can now see that in the learning phase there is still a tendency to cooperate following a temptation trial. This is markedly increased in the asymptotic period which may indicate an attempt to re-establish coordination after an anti-coordinated trial. Finally, examining the role of pellets received, we see that reward has a positive influence on cooperation, which increases from the learning to the asymptotic sessions,

In the isolated iSH condition rather than seeing an increased tendency to make the same choice as their opponent with learning, we instead see a slight decrease. Examining choice, we see a greater tendency after one trial into the past for rats to make the same choices. The history of temptation trials shows positive, but reduced, influence upon cooperation than in the standard condition. The number of pellets received shows a similar but slightly larger positive effect upon cooperation, particularly in the standard condition indicating that despite showing higher levels of defection overall, animals in this condition may be more influenced by outcomes than in the standard condition. This difference is likely due to the role played by social interactions in the standard condition that are absent in the isolated condition.

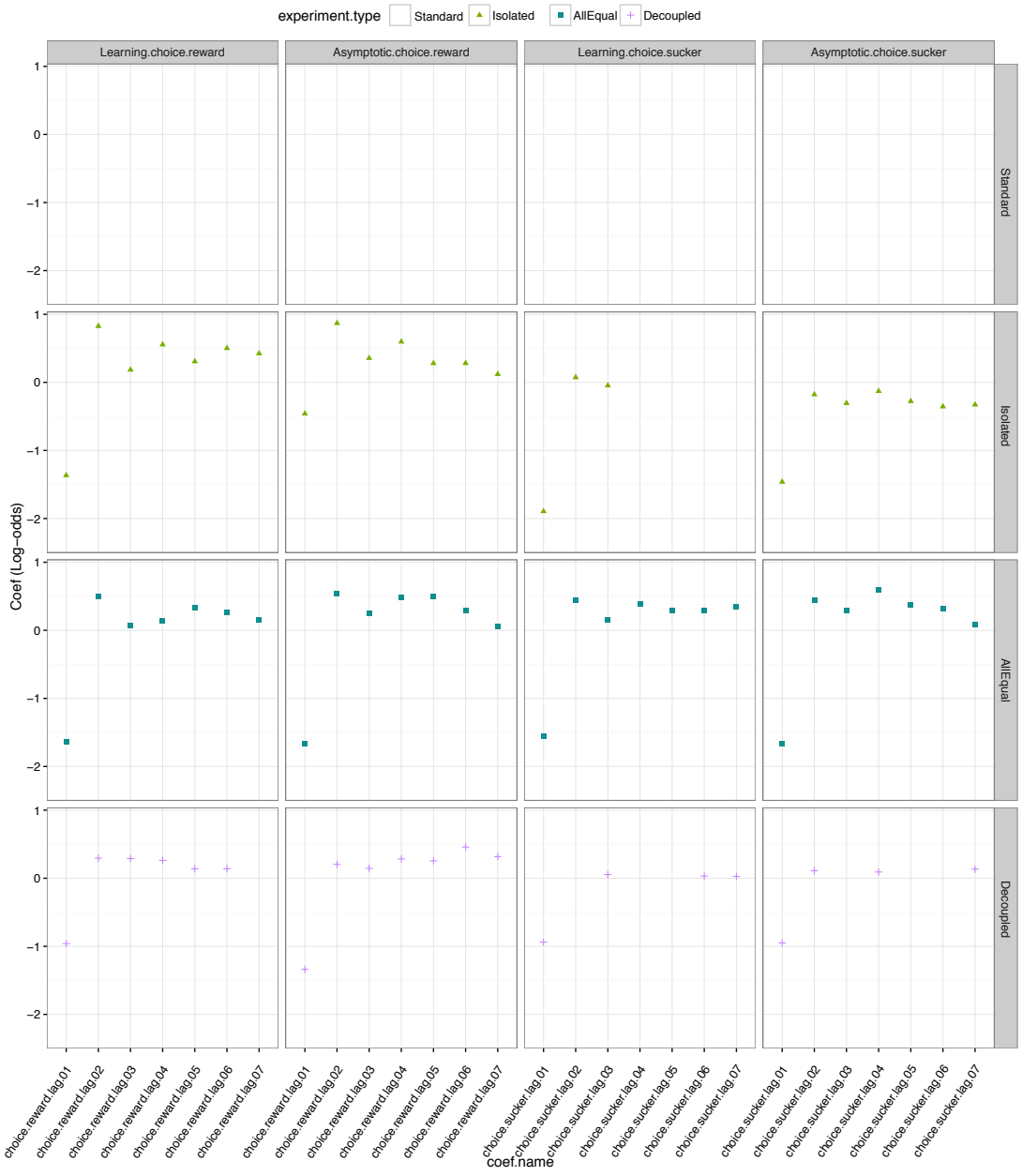


Figure 3.6.4 | GLM History of Cooperation outcomes (R, S) for learning and asymptotic periods

General linear model for Standard iSH (red, circles), isolated (green, triangles), all equal (turquoise, circles), and decoupled (purple, cross). Absent data points removed for no effect by L1 penalty.

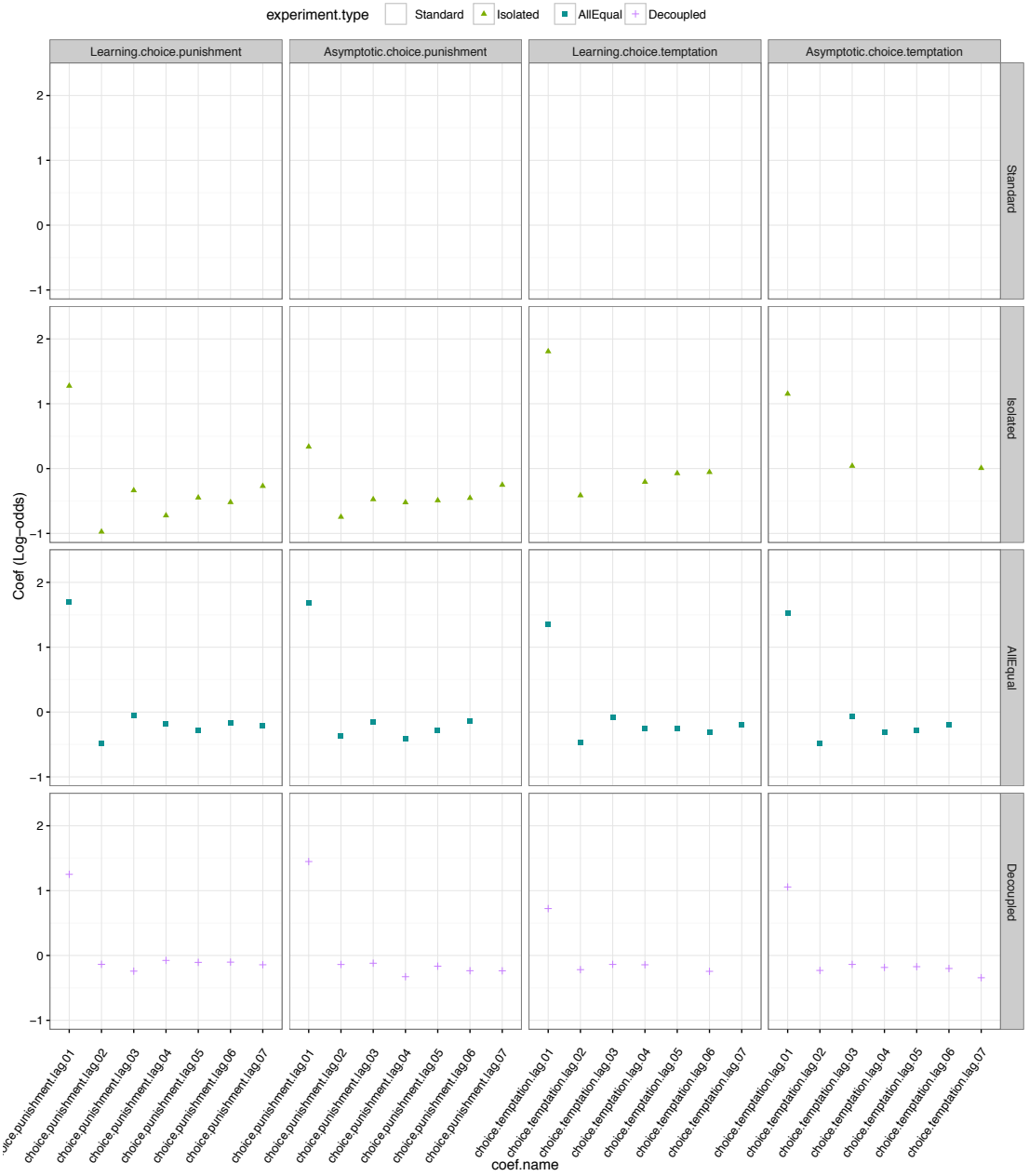


Figure 3.6.5 | GLM History of Defection outcomes (T, P) for learning and asymptotic periods

General linear model for Standard iSH (red, circles), isolated (green, triangles), all equal (turquoise, circles), and decoupled (purple, cross). Absent data points removed for no effect by L1 penalty.

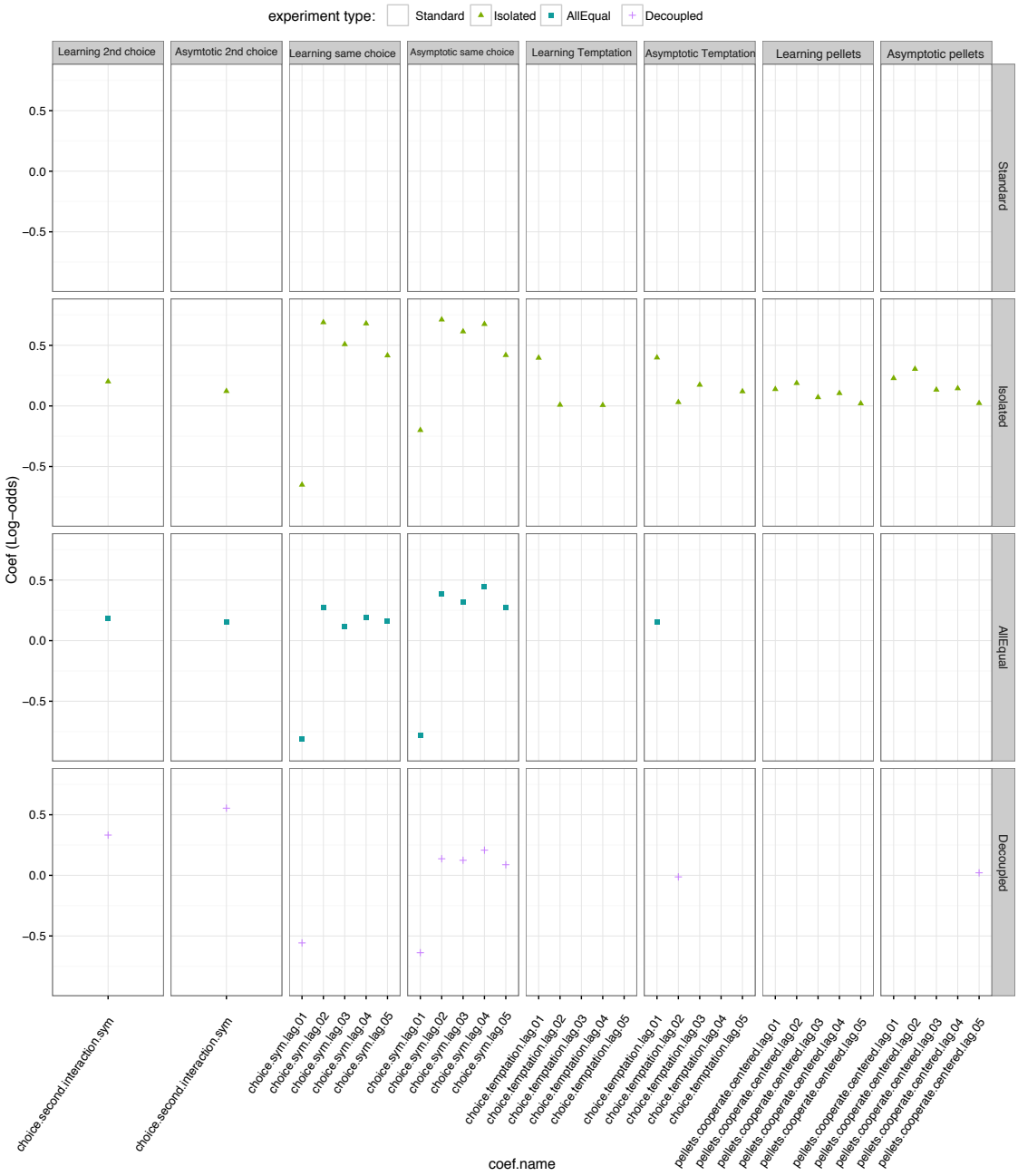


Figure 3.6.6 | GLM Full Model – second choice interaction, perseveration, history of choice, temptation and prior reward.
General linear model for Standard iSH (red, circles), isolated (green, triangles), all equal (turquoise, circles), and decoupled (purple, cross). Absent data points removed for no effect by L1 penalty.

3.6.6 Summary

Here we first explored the influence of the history or cooperate outcomes on future cooperation and separately, the history of defection outcomes on future cooperation. We then generated a more complete model that attempted to capture the most comprehensive collection of relevant influences upon choice behaviour. In the first two models, we can see that each model is consistent with previously observed alternation behaviour. In each condition in both learning and asymptotic periods, we see that both R and S outcomes predict defection in next trial, whereas in the second model T and P outcomes predict cooperation in the following trial. In the second model we included a 'switching kernel' that accounted for alternation and examined the influence of other's choice, pellets received and temptation upon future cooperation and the likelihood of making the same choice.

In the first model, we observed that the influence of prior S outcomes presents perhaps the most marked change as animals progress through the learning and asymptotic periods in the standard condition and as they transition to the isolated condition. Initially S trials have a marginal positive influence on cooperation, this increases substantially in the asymptotic period. This indicates that rats display a vastly increased tolerance for receiving the lowest outcome in later sessions. Following the removal of social information, this motif disappears almost entirely. In the asymptotic period of the isolated condition, S outcomes predict defection rather than cooperation. This indicates that social information is important for establishing and maintaining a tolerance to S outcomes. This is further supported by a marginal but consistent effect of prior S outcomes

predicting cooperation in both the learning and asymptotic period in the decoupled condition. In the isolated condition, particularly following the learning period, rats respond to outcomes in a more conventional manner; increasing the likelihood of cooperation following a positive R outcome and decreasing this likelihood following an S outcome.

The second model explored the effect of defect outcomes T and P on future cooperation, this model provides the one case where the outcomes were equivalent across all conditions. The first conclusion to take form is that P outcomes across conditions are in general quite similar with each predicting defection several trials into the past, both in the learning and asymptotic periods. T outcomes show a similar pattern, however, this effect disappears in the asymptotic period of both the standard and isolated condition but not the others. This indicates that regardless of social information, outcome structure and whether outcomes are in general contingent upon collective choices, the defection choice has similar, if not identical, influences on cooperation. This indicates that the differences we see across conditions, asides from the equal rewards condition, appear to be mediated by the outcomes received from cooperative choices.

The final, more complete model reveals some striking contrasts between the isolated and standard conditions. Here, we see an increase in cooperative coordination from the learning to asymptotic period in the standard condition, and the opposite in the isolated condition indicating that in the presence of social information, animals are learning to reciprocate cooperation, and in isolation, decrease this behaviour. We also see that in the absence of social information, animals increase in

perseveration over sessions, whereas there is only a marginal change with learning in the presence of social information.

Notably, the propensity of animals in the isolated conditions to cooperate appears to be similarly outcome driven as in the standard condition, and while each shows an increase with learning, this effect is greatest when social information is not present. This suggests that social learning strategies rather than purely outcome based learning may lead to increases in cooperation.

3.7 Discussion

In the experiments above, it has been demonstrated that freely behaving rats have the capacity and tendency to engage in mutual cooperation in an iSH task with access to ongoing social information. In so doing, while animals increased their received rewards, their behaviour deviated from economic expectations in a number of ways.

Firstly, animals did not behave in a manner that is consistent with the Nash equilibria of the Stag Hunt game. While they consistently coordinated their choices and cooperated at a rate greater than chance, they showed a strong preference to engage in alternation, being most likely to change their choice following mutual defection (P) and being more likely than chance to choose differently following mutual cooperation (R). The Nash equilibrium concept would instead predict that animals should perseverate in these choices, as they dictate that a unilateral change in choice would decrease their experienced outcomes. The primary explanation for this is likely to be, the by now well mentioned fact, that

animals in general do not often correspond precisely to the expectations of rational actors. However, alternation appeared to be a such a ubiquitous aspect of the animals' behaviour in all conditions, particularly in the equal rewards condition that it requires further explanation. In the work carried out by Ruth et al (2016), they found that cooperation was more likely when it was signaled by rats pressing levers, rather than withholding lever presses[53]. Similarly here, a strong possibility for this behaviour may be a function of the particular design of the assay itself. Early work examining the behaviour of rats in mazes demonstrated that even in the absence of rewards, in T-mazes, rats show a strong spontaneous preference for alternation[55].

This may seem a trivial point at which to begin the discussion of the prior experiments. However, it may add an important element to the dilemma each animal faced. In principle the easiest way to coordinate choices in this task is for each animal to always choose the same side, which is supported by the predictions of the Nash equilibrium. However, in chapter 2 we saw that rats coordinated more rapidly with a pseudo random stooge, that therefore largely alternated, than with a stooge that always cooperated. While they did eventually engage in robust and consistent cooperation with an all cooperate stooge, it took several sessions to establish. Alternation was also at its highest when it had no effect upon received outcomes in the all-equal condition indicating that this behaviour was not driven by an attempt to increase rewards. If this behaviour is indeed spontaneous, and, as there would be no reason to expect otherwise, present in both members of a dyad, and as experiments in chapter 2 suggest, is something that requires some learning to overcome, then it places a further constraint upon coordination and thus

cooperation. It makes the simple strategy of perseveration in one Nash equilibrium or the other less likely.

While this may seem more like a 'bug' than a 'feature' in this task, one could argue that it does add to the richness of the dilemma. Similar to other more cognitive limitations, such as temporal discounting[21], it provides a hurdle to cooperation. However, in the freely behaving case it also becomes an important component of the animals' interaction. Rats that are attempting to maximize reward through increasing cooperation, must overcome this tendency to alternate by choosing to increasingly perseverate in cooperation. For this to increase reward, it must also be done in a coordinated manner so that R rather than S outcomes are more frequently achieved. The tendency for each animal to coordinate their alternation is precisely what is measured by coordinated switching, the conditional probability of making the same choice given that either rat has chosen differently than in the prior trial. In the standard condition, the proportion of coordinated switching trials does not increase notably over sessions, however the proportion of mutual cooperation does. We also see over sessions, an increased likelihood of longer bouts of cooperation in this condition. This suggests that animals do not learn to improve their ability to coordinate their alternation, but rather learn to collaboratively bias this alternation to increasingly favour the cooperate arm.

High levels of alternation add an extra layer of complexity to coordinating choices and therefore, in dyads maintaining high levels of cooperation. However, in the standard iSH condition animals show a consistent capacity to coordinate in alternation. This capacity for coordinated switching was only observed when animals had access to

social information. The only case where it fell to chance levels was when previously cooperating animals were transitioned to the isolated condition. This failure to coordinate alternation was accompanied by a sharp shift in preference from mutual cooperation to mutual defection. Animals, it appears, require social information to couple their alternation behaviour effectively. When they fail to do so, they increasingly bias their alternation to favor defection, which provided constant reward.

Animals in the decoupled condition developed no preference for cooperation or defection, despite the presence of conspecific and the same likelihood of receiving R and S outcomes on average. This indicates that high levels of mutual cooperation observed in the standard condition was not meditated simply by the presence of conspecific, but rather by the contingency between their opponents' choice, their own and the outcome.. This strongly suggests that an average increase in reward is not sufficient to support cooperation. Rather that each rat's opponent's choices reliably predicted the available outcomes in a given trial was also required for high levels of cooperation.

Echoing the account in the introduction, anecdotally, a similar 'dance' between animals was particularly noticeable in the standard iSH condition. While not as dramatic or as long in duration, animals appeared to track each other's movements prior to choice. Animals could be seen to double back when they started to move to one choice arm but their opponent had already moved to choose another. Video tracking of these animals was conducted and will be analysed to quantify these effects. This further suggests that trial-by-trial coordination of choices depended upon the use of ongoing social information and this information had to provide a

meaningful indicator of outcome to increase cooperation. This adds a further departure from economic assumptions, that choice is based purely on received prior rewards.

One notable difference between the standard condition and each of the other condition is that it is the only case where the likelihood of switching is lowest following an S outcome. It was also the only condition where this was lower than chance, and the only case where it decreases over sessions. This tendency to perseverate is very much like the 'forgiveness' seen by Stephens *et al* (2002). That rats in this task are learning to behave in a manner consistent with this concept of 'forgiveness' was further supported by the result that S outcomes more than a single trial into the past predicted future cooperation in the asymptotic sessions.

Perhaps the simplest explanation for this tolerance of low S outcomes would be that animals develop a habitual preference for the cooperate arm. This seems somewhat unlikely as a preference for alternation is maintained, though the pattern of alternation they engage in could also be habitual in nature. The results from the reversal experiment argue against this position as dyads react rapidly to the reversal of the cooperate arm, quickly re-establishing a preference for cooperation on the novel arm.

Another potentially important aspect of Stephens et al's (2002) findings was that this tolerance for S outcome was revealed only when they controlled for temporal discounting and used reciprocal stooges. One important difference between the Stag Hunt and Prisoners Dilemma is that

the Stag Hunt reduces the requirements for control of temporal discounting[35]. This is simply a result of having R provide the greatest outcome rather than T, meaning that the net reward of mutual cooperation is received directly rather than being accrued over multiple trials.

It is therefore possible that reducing the role of temporal discounting would be sufficient to support this 'forgiveness' behaviour rather than there being a requirement for use of social information. While temporal discounting may play a role in this behaviour, it cannot be considered sufficient as it was not observed in the isolated condition where animals were most likely to switch following an S outcome. Also, sucker outcomes more than one trial into the past had a negative influence on future cooperation when social information was absent.

The data in this chapter demonstrates that rats can indeed hunt stag. However, they are only able to do so in the presence of conspecific and when the social information this provided represented the contingency between collective action and outcome. This permitted freely behaving animals to coordinate their alternation and bias it in the direction of the payoff dominant equilibrium.

General Discussion

"You can add up the parts
but you won't have the sum"
Anthem, Leonard Cohen

4.0 Summary of empirical findings

In this thesis, we focused upon the contributions that social and economic information have in supporting cooperative behaviour. More specifically, we posed a particular social dilemma, the Stag Hunt, to pairs of freely behaving rats and asked whether they were first able to coordinate their choices, and then whether they were able to do so cooperatively. We began by detailing the design of the assay that was used. We then determined via simple discrimination experiments a set of rewards that were both discriminable and desirable to rats.

By the use of experimentally controlled stooge animals, we then proceeded to determine if animals could combine social and reward information in a manner consistent with the predictions of the Stag Hunt game, in this case, to coordinate. We approached this by determining whether they were capable of associating their own choice, the position of a stooge animal, and outcome to optimize reward. We first used a stooge whose position on the cooperate or defect arm was determined from a pseudo random distribution. This ensured that only trial by trial social information could be used by the agent animal to guide their choices. Agent animals demonstrated a capacity to follow the stooge animal on the majority of trials, indicating that they were capable of making this association. In this case, agent rats were able to coordinate from the first session, indicating a preference for proximity with the stooge animal that did not require learning.

We then determined the capacity of agent rats to behave in a manner consistent with the payoff dominant Nash equilibrium by pairing

them with an all-cooperate stooge. Agent rats, after several sessions, showed that they were indeed able to behave in a manner consistent with the Nash equilibria by almost exclusively cooperating with the stooge animal. To distinguish between whether this was merely a result of preference for proximity to the stooge animal or whether this preference for coordination was motivated towards increasing rewards, we then exposed agent rats to a modified version of the Stag Hunt payoff matrix that rewarded anti-coordination rather than coordination in the context of pseudo random stooge. Agents rats produced an intermediate response in this condition, decreasing considerably their tendency to coordinate, but rather than shifting to a more rewarding anti-coordination behaviour, both naïve agent rats and those that had previously experienced that standard payoff matrix, instead rapidly increased their tendency to engage in defection. This indicated that animals under the standard payoff matrix were not simply following the stooge animal, but rather were coordinating to increase their reward. However, it remains unclear if animals failed to anti-coordinate because they were unable to associate the choice not made by the stooge with increased reward, leading to a preference for the safer defect choice or if there was a conflict between the value of the reward and proximity to the stooge animal that resulted in a preference for defection.

Having established that agent rats were capable of behaving in a manner consistent with an understanding of the association between collective action and reward, we then proceeded to examine the behaviour of freely behaving animals. We examined the behaviour of freely behaving dyads under five separate conditions. The equal rewards condition provided a series of ad-hoc baselines that provided insight into how animals responded to the assay itself and to the presence of a conspecific.

This revealed, primarily, that in the absence of economic incentives to do so, animals would consistently and moderately coordinate, however, not at a rate that would have removed the risk of cooperation. They also showed a strong tendency to alternate between one choice arm of the assay and the other. Three conditions used the standard iSH payoff matrix, in one case animals had complete access to social information, then following fourteen sessions in this condition, we removed social information and observed the effects upon behaviour in a subset of these animals. Similarly, we also conducted a reversal condition where following completion of fourteen sessions, a further subset of animals was then exposed to the standard iSH payoff structure. In this reversal condition, dyads quickly re-established similar levels of mutual cooperation to that prior to the reversal, demonstrating notable behavioural flexibility in the face of changing outcomes. This also excluded the possibility that cooperative behaviour in the standard iSH condition was habitual.

Given access to social information and the standard iSH payoff matrix, freely behaving dyads demonstrated a tendency to alternate, making high rates of mutual cooperation a more complex problem. Nonetheless they also showed a consistent capacity to coordinate this alternation and within six sessions demonstrated a robust capacity to mutually cooperate. Rather than engaging in behaviour consistent with standard application of the Nash equilibrium-like behavior, animals instead progressively biased their alternation towards cooperation in a coordinated manner. Removal of social information resulted in a rapid decrease in the ability of dyads to coordinate their alternation. This was accompanied by a rapid decrease in mutual cooperation, and an increasing preference for mutual and unilateral defection. This demonstrated that

ongoing social information was fundamental for maintaining both coordinated alternation and mutual cooperation.

In a final condition we provided the same likelihood of receiving R and S outcomes for cooperation as in the Standard iSH condition regardless of the choice of each animal's opponent. This provided animals with greater reward *on average* for cooperating than defecting. In this condition where social information was present but did not provide information about reward, animals continued to coordinate their alternation. However, they did not develop a preference for unilateral or coordinated cooperation meaning they continued to coordinate in a manner that reduced their overall reward rate.

4.1 Social information, cooperation and forgiveness.

Freely behaving animals only showed consistent cooperation in the standard iSH condition where the relationship between collective choices and reward followed the standard iSH payoff matrix and each animal had complete access to social information. This adds to several existing examples in rats[32,33], capuchins[34] and chimpanzees[35] where cooperation in social dilemma tasks was increased by the presence of social information. Similarly, Marquez *et al* (2015) demonstrated that rats were capable of responding to the expressed preferences of other animals. Coordinated switching provided a powerful measure of the capacity of dyads to respond in a trial-by-trial manner to the behaviour of their opponent. The consistent tendency to alternate across almost all conditions, later sessions with an all-cooperate stooge being the only observed exception, places a premium upon social information. The tendency to alternate makes achieving high rates of mutual cooperation all the more a collaborative problem. It

requires each animal to synchronize their choices with their opponent, and bias those choices, in coordination increasingly towards cooperation. Having access to the ongoing flow of social information provides each animal with evidence, albeit imperfect or ambiguous evidence, as to the choice of their opponent. It also provides them with a means of collaborating, of collectively orchestrating their choices on a trial-by-trial basis.

The presence of social information, while important for dyads to coordinate their alternation, may also have an influence in the 'reward phase' of a given trial, when each animal receives rewards for their respective choices. Schuster *et al* (2001) demonstrated that following sessions in which pairs of rats had coordinated in a social context, they increased their consumption of a sucrose solution over that of a non-social context. Marquez *et al* (2015) also showed that while indications of a side preference by the recipient rat supported pro-social choice by an agent animal, it only did so when the recipient of that pro-social choice received increased reward as a result. This indicates that in the very same assay used here, rats are also sensitive to the reward received by the other animal. It is also possible that the presence of a conspecific during reward delivery in R trials increases the value of that outcome beyond the simple economic difference.

Animals in the standard iSH condition also demonstrated increased tolerance for S outcomes. In the asymptotic phase, animals were most likely to repeat their choice in subsequent trials following an S outcome. This was supported by the GLM analysis, which indicated that S outcomes further into the past also predicted cooperation. This behaviour was also

noted by Stephens et al (2002) during an iPD task with a reciprocal stooge in a condition where the requirement to minimize temporal discounting was reduced. Similar tolerance for S outcomes was found when zebra finches engaged in an iPD task with a pair bonded animal compared to a novel opposite sex conspecific[36]. In each case this tolerance is termed 'forgiveness' in the strategic sense, in animals that are willing to forgive cheating on the part of their partner and continue to cooperate despite having experienced exploitation. Both explanations, a relaxing of the requirements to engage in temporal discounting and the presence of strong social bonds provide plausible explanations for the forgiveness behaviour observed here.

The Stag Hunt relaxes the requirement to minimize temporal discounting by making the greatest reward contingent upon mutual cooperation (R) and therefore immediately available. Animals in these experiments were in all cases cage mates, which may also represent a similar, though clearly distinct from a mating pair, long-term social bond. In the case of the zebra finches, increased forgiveness was displayed towards a pair-bonded partner and this was speculated to be the result of an increased likelihood of future interactions, which favors increased reciprocity. This suggests that examining the behaviour of animals in their home cage following more or less cooperative sessions, or perhaps standard and isolated iSH sessions may provide further insights into how interactions outside of the current experimental context may influence and be affected by cooperation and social information.

One concern with the display of tolerance seen in the standard iSH condition is that a preference for cooperation, one that was more habitual

in character, would produce behaviour that was very similar. This we initially addressed via the reversal condition where it was observed that animals rapidly re-established both individual and mutual cooperation after the cooperate and defect choice arms were reversed. A habitual explanation also was not supported by the GLM analysis which demonstrated that, despite this tolerance, animals in this condition were still using the history of rewards to guide their decisions, albeit in a forgiving manner.

By changing the relationship status of animals, comparing non-familiar animals or by orchestrating negative interactions between members of a dyad, one may be able to distinguish between the role of temporal discounting and the role of familiarity as potential causes for this forgiving behaviour. Another means of approaching this would be to modify the payoff matrix so that the reward received for mutual cooperation following a forgiving trial for the previously forgiving animal was lower. This would provide a more precise means of examining what determines an animal's adherence to this behaviour.

4.2 The meaning of social information

A decoupled condition was employed to replicate the reward statistics that dyads encountered during the asymptotic phase of the standard condition, but to remove the contingency of these rewards upon collective choice. In so doing, we sought to determine if high levels of cooperation required a high likelihood of receiving an R outcome or the presence of conspecific, both of which were absent in the isolated condition, or if the contingency between collective action and outcome mattered. We reasoned that the meaning of social information was also

important. Despite animals having the opportunity to reliably increase their reward on average through cooperation, dyads did not display a preference for cooperation or defection in this condition. Animals in this condition were more likely to cooperate than in the isolated condition, indicating the presence of conspecific and reward statistics that support cooperation were not sufficient to drive the levels of cooperation seen in the asymptotic phase of the standard iSH condition. This suggests, on one hand, that animals are using social information in a trial-by-trial manner to support cooperative choices. On the other hand, despite incurring a cost to do so, each animal in this condition continued to engage in a high proportion of coordinated alternation. That each animal had a preference for coordinating with its partner over cooperation may suggest the presence of a conspecific may have actually interfered with reward guided decision making. Each animal may have been attempting to infer the underlying distribution that determined their likelihood of receiving rewards, and therefore, in the face of uncertainty, continuing to coordinate with their opponent.

This possibility, however, still requires that animals used social information to guide their choices on a trial-by-trial basis. This is similar to the anti-coordinate iSH stooge condition, where animals were more likely to cooperate when the stooge cooperated than when it defected despite incurring a cost to do so. In each case social information may have functioned as a distractor. This data was only collected in the final weeks of this thesis. As a result, we have not analyzed the extent to which animals change their choice contingent on reward and have only done so based on collective choice. Subsequent analysis of this data will examine the likelihood of switching following reward received, in the hope of

further disentangling the role of outcome and social information. Another means of dissecting the potential role the presence of a conspecific has upon the animals' propensity to cooperate, and whether the presence of a conspecific acts as a distractor in this case would be to conduct the same condition, but in isolation. One could then determine whether animals develop a preference for cooperation given the same outcome statistics.

The tendency of animals in the decoupled condition to continue to coordinate their alternation with their opponent in the absence of bias to cooperation raises an interesting possibility. It may be that coordination may be more important to the animals than the moderate reward increase offered by increasing unilateral cooperation. This is not to say that animals are simply following each other, as the stooge experiments demonstrated. Coordinated switching is higher in the decoupled condition than in the equal rewards condition, indicating that rewards are playing an important role in supporting coordinated choices. It may be that coordination becomes more valuable in the context of unpredictable reward, either because each member of the dyad considers its opponent a salient cue upon which to predict reward, even when it is not. Or it may be that in the face of uncertainty, coordinating their choices, actively collaborating to overcome the uncertainty is in itself more valuable to each animal than the small increase in reward offered by unilateral cooperation.

4.3 Social Competence and Cooperation

In the introduction, it was proposed that there was the possibility to ground game theoretical approaches to cooperation within the framework of social competence proposed by Oliveira and Taborsky. This potential

approach to cooperation, one more focused on the interaction between animals and their use of social information, had three conditions. First, each animal must honestly signal their intentions, and in so doing be more predictable to a conspecific with which they interact. Second, there must be an attempt to interpret the signals of the conspecific they interact with. Third, their response must increase the likelihood of benefits, or limit the cost of the interaction for both themselves and the interacting conspecific. These three requirements seem to be met by the behaviour of animals in the iSH task. In the standard iSH condition, animals appear to be using trial-by-trial, proximate social information to guide their choices, as we have seen in high rates of coordinated alternation. The Stag Hunt payoff matrix makes it beneficial for each animal to be predictable to its opponent and to predict that opponent's choice as coordination leads to increased rewards. This is particularly true when animals have access to social information throughout the decision-making phase in each trial. Animals in the standard iSH condition robustly engaged in cooperation, satisfying the third condition; increasing their own potential benefit and that available to their opponent. Further, as cooperation decreased in isolation, it suggests that key aspects of social competence, rather than more instrumental mechanisms were being employed to support cooperation.

The distinction between reward learning and learning to use social information, in the light of social competence theory, thus appears to integrate many of our observations across experimental conditions. Notably, animals in the standard iSH condition coordinated at a rate higher than that of agents with a pseudo random stooge under the standard iSH payoff matrix. This is surprising as the stooge animals moved to either the cooperate or defect choice arm first, providing agent animals with clear

information on what choice they should make before they made it. In the standard iSH condition, animals had simultaneous access to the choice arms, which on the face of it should make coordination more complex. However, in the standard iSH case, both animals were unconstrained not just in their choices but also in the fidelity of the social signals they displayed to make their choice. It seems that in the stooge case, we provided agent rats with 'ideal' information with which to cooperate, and yet, the issues mentioned in that chapter notwithstanding, they seem less capable of coordinating on average. It may be that stooge animals as their 'choices' were forced, did not display clearly interpretable social signals.

By controlling the movements of animals, similarly to Viana *et al's* iPD work, the stooge animal may have served as more of a cue than a collaborator. This may have emphasized the instrumental aspect of the task rather than one that examined social competence. Similarly, when social information was removed, animals' choices became more typically instrumental in that they were less likely to cooperate following an S outcome than in the standard condition.

An important aspect of social competence theory is to examine behaviour in multiple ethologically relevant social contexts. This was not performed here, and therefore one cannot make the claim that the observed behaviour is based on a more general behavioural flexibility. Data presented here that might speak to this may be the reversal condition. While not a different context, it does indicate that animals are able to flexibly and collectively regulate their behaviour in the face of changing conditions. However, examination of home-cage behaviour or in other social assays such as the coordinated shuttling task[32] or other social

foraging tasks would be required to demonstrate that the coordination and cooperation behaviour found here was representative of a more global social competence.

A further way of examining whether social competence provides useful insights to the work conducted here would be to conduct closer analysis of individual variation. Some animals and dyads cooperated more than others. Thus far, the behaviour of dyads has been approached on the population level. It may be that there are subtle differences in animals and dyads that cooperate more or less. This would facilitate the examination of what Taborksy and Oliveira term behavioural reaction norms (BRN's), which refer to the set of behavioral phenotypes produced by a single individual in a particular context. For example, there may be subtle strategic differences between animals such as the propensity to 'forgive' S outcomes or to control their alternation in a social context. Comparing the variation seen in this task, i.e, greater levels of forgiveness, and examining it in a novel experimental context would also be an effective way of determining if this particular tendency is part of a more instrumental approach to increasing reward, or is indicative of a more general social disposition.

Perhaps one of the more interesting insights that a social competence approach would bring to these series of experiments is that the social experience of animals should play an important role in their behaviour. Comparing animals that were developmentally isolated, for example, with those that were not may precisely reveal the role of social competence in this type of cooperative behaviour. Or, as this may also influence instrumental as well as social capacities, individually housed vs.

those housed in larger groups may also provide a fruitful comparison. One hypothesis one could draw from this is that animals that were raised in isolation or held in less socially enriched contexts would demonstrate a decreased capacity to emit or follow the social signals that support high levels of coordinated alternation.

4.4 Cooperation, cooperation *or* cooperation?

The introduction to this thesis began by pointing out that cooperation is notoriously tricky to define, in part because there are many approaches to the issue. Cooperation has been used throughout this thesis to refer to very different proximate situations. It referred to one of two available choices in stooge experiments, though the stooge animal received no benefit from a cooperative choice, as its reward was always a single pellet. It was used in the isolated iSH condition, when animals had no evidence that their choices influenced the outcomes that their partner could receive. In the standard iSH case, animals that cooperated did provide their opponent with the possibility of increased reward, however whether individual rats were sensitive to this fact was not conclusively demonstrated. It was also used in the decoupled condition, where outcomes for each animal were independent. Nor do any of these animals' choices have any influence upon the fitness of the rats involved. The simplest answer that is consistent for all these conditions is that cooperation was an operational definition, an arbitrary distinction between each choice arm for each animal or dyad. However, this definition reduces to the instrumental, self-regarding definition used in game theory. Another way to approach this is to consider the standard iSH, the central experiment and the other conditions, those that involved stooges and the other conditions with freely behaving dyads, as a series of control

experiments. In the standard iSH experiment, animals developed a preference for mutual cooperation but for this choice to be valuable animals had to work together to bias their alternation towards cooperation, and do so in a coordinated manner. It is unclear whether each rat had any knowledge that their choice to cooperate provided the possibility of an extra reward pellet and removed the possibility of their opponent getting 0 pellets if they engaged in unilateral cooperation.

And yet both members of each dyad were able to overcome their tendency to alternate sufficiently, overcome the cost of being suckered, and coordinate their choices to an extent that each benefited more from the interaction than if they had simply defected. In light of this, the definition drawn from social competence, to predict, be predictable and to do so collectively to increase mutual benefit does seem like the most pertinent functional definition of such cooperative behaviour. Or more simply put, cooperation here describes a capacity for collaborative interdependence under risk.

Previously, it was proposed that the particular structure of the payoffs used in this version of the Stag Hunt lay somewhere between a typical mutualism, where cooperation is always the better option, and reciprocal altruism, where the benefits of cooperation are only received after a delay. Mutualisms are considered to be common in animal societies while reciprocal altruism considerably, less so. Tomasello and others have postulated that mutualisms serve as an evolutionary precursor for more complex and cognitively demanding forms of cooperation. He proposes that the opportunity to engage in mutualistic interactions places a premium upon the interdependent collaboration of unrelated actors and establishes

the motivational and cognitive foundations that give rise to the types of socially competent cooperation we have described here. Indeed, the work mentioned in the introduction to chapter 3 that compared human infants and chimpanzees found that chimpanzees cooperated at rates very similar to rats in this experiment[35]. The data here, in combination with the work of many other groups, is increasingly demonstrating rats are sophisticated social animals. Here, we have shown that they are capable of signaling their intentions, integrating social signals and economic information, overcoming risk, and engaging in sophisticated socially competent cooperation.

4.5 Future Directions

4.5.1 Video Analysis

The most pressing aspect of this work and that which will be embarked upon directly is analysis of video data from each condition. In all of the freely behaving conditions, a color based tracking system that allows for independent tracking of both the centroid of each animal's body as well as their head direction has been implemented. This provides a potentially extremely rich source of data that will hopefully get closer to precisely how rats in this task are making use of moment-to-moment social information which requires some speculative comment$_8$.

Thus far, we have approached this only at a trial-by-trial resolution, focusing on animals' ability to coordinate their alternation. As animals

[8] The range of potentially interesting analyses that could be performed on this data is quite large, and as it is speculative, I will provide only a few of the many possibilities.

learned to progressively bias their coordination towards cooperative choices in the standard condition, and as this seems distinctly collaborative in nature, examining the changes in the movement patterns of individual animals and the relationship between them within each dyad may provide some indications of how this collaboration takes place. For example, in the standard condition, the choice to cooperate is risky, while choosing to 'cooperate' in the all equal condition is not. Again the all equal condition may provide a useful ad-hoc baseline against which these differences can be compared. We already observed that in the all equal condition, there were more pairs where one animal chose first more frequently than in the standard condition. One could hypothesize that this is because when cooperation is risky, animals wait to accumulate more evidence about the choice their opponent will make. This information may be signaled posturally, for example, via head direction, or positioning of the body closer to either the cooperate or defect arm of the assay. One way of approaching this would be to see if cooperative choices in the standard condition coincide with increased correlation in the head direction of each member of dyad than in the in the equal rewards condition. By categorizing and stereotyping the individual animal and inter-individual, within dyad, movement patterns for cooperation, defection, coordinated and anti-coordinated choices, one may be able to reveal what type of social signals each animal is sending and receiving[9] and perhaps a measure of across dyad sensitivity to these signals. This may provide the most comprehensive way of approaching social competence in the context of this task.

[9] Vocalizations of each animal were recorded for many sessions, which presented a promising approach to the same question. However animals vocalized so rarely that this was discontinued, further emphasizing the importance of other social cues.

Analysis of the animals during the reward phase of each trial may also shed light on differences in behavior between R and P trials. Animals may interact more, or simply differently, following R trials. Similarly, 'forgiving' trials may involve overt signaling on the part of the forgiving animal, which may encourage their opponent to cooperate and help explain the particular pattern of switching behaviours observed in the standard iSH condition. Contrasting across conditions, for example, comparing responses in the standard iSH condition with the all equal and isolated condition following S outcomes, may provide useful insights into communicative behaviours.

Possibly the more conservative, but not less important, insight that can be drawn from this video data is to gain a better measure of which animal chooses first. Cooperating first incurs the risk of receiving an S outcome. Despite the importance of the order of cooperative choices, choice order has not received much attention throughout this thesis. This is simply a result of the fact that choice time was recorded when each animal reached the furthest point of each choice arm. While this is the point at which choice in a given trial is irreversible, it does not exclude changes in the speed of each animal as they approach the reward site. Anecdotally, while animals do double back and revise their choice on a subset of anti-coordinated trials, they have not been observed to do so on coordinated trials. This may suggest that animals perceive their opponent as having signaled their choice some time before reaching the reward site. Having a more precise measure of the time course of each collective choice will likely shed more light on how dyads manage the risk of cooperating first.

4.5.2 The role of learning

In chapter 3, we demonstrated that animals learned to increase mutual cooperation in the Standard iSH condition and upon removal of social information, the same animals learned to adopt a pattern of behaviour characterized by increased defection. While we have compared differences between learning and asymptotic period, we have not further characterized the learning process itself. One way of approaching this is to apply a reinforcement learning approach. Macy and Flache (2002) have established a reinforcement learning approach to social dilemma tasks. Such an approach has not been previously applied to freely behaving animals. Such an approach can shed light on concealed cognitive variables, such as changes in reward expectation from trial to trial. We have already begun to apply this model to the behaviour of each rat in a dyad with some early success. Applying such a reinforcement learning model may provide a novel approach to animal behaviour in a social dilemma task, it may also provide a novel source of information against which to compare video data.

4.6 Final Remarks

It is perhaps the way of all theses that they feel unfinished. That they present more questions than they answer, and perhaps that is their purpose, to describe the inevitably incomplete with some finality. At this point, I wish to thank again those that have contributed directly and practically to this work. Marta for her careful support, and for endless patience. Eric Dewitt for his admittedly less constant, but fundamentally important, and at times tireless support and guidance. Christina Marquez for her insight and collaboration. And to Alexandra Silva for collecting a significant portion of the data in Chapters 2 and 3 and in so doing, sharing

many of the frustrations I have felt. It is pleasing, in the simplest sense, that this discussion of cooperation emerged from a deeply a collaborative and at times entirely altruistic effort on the part of those who contributed to it. Sometimes people behave as though science were a mutualism, if enough did, it would be.

While this project is far from finished, the data detailed in these chapters and the ideas discussed, I hope, present a novel, ever so slight, incremental step in understanding cooperation. The integration of social competence and game theoretical approaches proposed here, albeit in its embryonic form, present a promising approach to understanding the enduring mystery of cooperation.

References

1. Taborsky, B. & Oliveira, R. F. Social competence: an evolutionary approach. *Trends Ecol. Evol.* **27,** 679–88 (2012).
2. Nowak, M. a. Five rules for the evolution of cooperation. *Science* **314,** 1560–1563 (2006).
3. Noë, R. Cooperation experiments: coordination through communication versus acting apart together. *Anim. Behav.* **71,** 1–18 (2006).
4. Camerer, C. F. & Fehr, E. When does 'economic an' dominate social behavior? *Science (80-.).* **311,** 47 (2006).
5. Camerer, C. F. *Behavioral Game Theory : Experiments in Strategic Interaction*. (Princeton University Press, 2011).
6. Hake, D. F. & Vukelich, R. A classification and review of cooperation procedures. *J. Exp. Anal. Behav.* **18,** 333–43 (1972).
7. Stevens, J. R. & Hauser, M. D. Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* **8,** 60–65 (2004).
8. Clements, K. C. & Stephens, D. W. Testing models of non-kin cooperation: mutualism and the Prisoner's Dilemma. *Anim. Behav.* **50,** 527–535 (1995).
9. Walker, M. M. On a Wing and a Vector: a Model for Magnetic Navigation by Homing Pigeons. *J. Theor. Biol.* **192,** 341–349 (1998).
10. Wojcik, P. T. *et al*. Single Scale for Odor Intensity in Rat Olfaction. *Curr. Biol.* **24,** 568–573 (2014).

11. Kepecs, A. The Sniff as a Unit of Olfactory Processing. *Chem. Senses* **31,** 167–179 (2005).
12. Wesson, D. W. Sniffing behavior communicates social hierarchy. *Curr. Biol.* **23,** 575–580 (2013).
13. King, R. S. & Newmark, P. A. The cell biology of regeneration. *J. Cell Biol.* **196,** 553–562 (2012).
14. Heyes, C. *et al.* Not-so-social learning strategies. *Proc. Biol. Sci.* **282,** 4–14 (2015).
15. Schuster, R. & Perelberg, A. Why cooperate?: An economic perspective is not enough. *Behav. Processes* **66,** 261–277 (2004).
16. Scott-Phillips, T. C., Dickins, T. E. & West, S. A. Evolutionary Theory and the Ultimate-Proximate Distinction in the Human Behavioral Sciences. *Perspect. Psychol. Sci.* **6,** 38–47 (2011).
17. Tinbergen, N. On aims and methods of Ethology. *Z. Tierpsychol.* **20,** 410–433 (2010).
18. Clutton-Brock, T. *et al.* Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* **296,** 69–72 (2002).
19. Packer, C. & Ruttan, L. The evolution of cooperative hunting. *Am. Nat.* (1988). at <http://www.jstor.org/stable/2461865>
20. FLACHE, A. & MACY, M. W. Stochastic Collusion and the Power Law of Learning: A General Reinforcement Learning Model of Cooperation. *J. Conflict Resolut.* **46,** 629–653 (2002).
21. Stephens, D. W., McLinn, C. M. & Stevens, J. R. Discounting and reciprocity in an Iterated Prisoner's Dilemma. *Science*

298, 2216–2218 (2002).

22. Lombard0, M. P. Tree Swallows and TIT FOR TAT.
23. Milinski, M., Pfluger, D., K□lling, D. & Kettler, R. Do sticklebacks cooperate repeatedly in reciprocal pairs? *Behav. Ecol. Sociobiol.* **27,** 17–21 (1990).
24. Hart, B. L. & Hart, L. A. Reciprocal allogrooming in impala, Aepyceros melampus. *Anim. Behav.* **44,** 1073–1083 (1992).
25. Hauser, M. D., Chen, M. K., Chen, F. & Chuang, E. Give unto others: genetically unrelated cotton-top tamarin monkeys preferentially give food to those who altruistically give food back. (2003). doi:10.1098/rspb.2003.2509
26. Olendorf, R., Getty, T. & Scribner, K. Cooperative nest defence in red-winged blackbirds: reciprocal altruism, kinship or by-product mutualism? *Proc. Biol. Sci.* **271,** 177–82 (2004).
27. Krams, I., Krama, T., Igaune, K. & Mänd, R. Experimental evidence of reciprocal altruism in the pied flycatcher. *Behav. Ecol. Sociobiol.* **62,** 599–605 (2008).
28. Viana, D. S., Gordo, I., Sucena, É. & Moita, M. A. P. Cognitive and Motivational Requirements for the Emergence of Cooperation in a Rat Social Game. *PLoS One* **5,** e8483 (2010).
29. Green, L., Price, P. C. & Hamburger, M. E. Prisoner's dilemma and the pigeon: Control by immediate consequences. *J. Exp. Anal. Behav.* **64,** 1–17 (1995).
30. Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement,. *J. Exp. Anal. Behav.* **4,** 267 (1961).
31. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Matching

behavior and the representation of value in the parietal cortex. *Science* **304,** 1782–7 (2004).

32. Schuster, R. COOPERATIVE COORDINATION Experiments with an Animal Model. *Hum. Nat.* **13,** 47–83 (2001).
33. GARDNER, R. M., CORBIN, T. L., BELTRAMO, J. S. & NICKELL, G. S. THE PRISONER'S DILEMMA GAME AND COOPERATION IN THE RAT. *Psychol. Rep.* **55,** 687–696 (1984).
34. Mendres, K. A. & De Waal, F. B. M. Capuchins do cooperate: the advantage of an intuitive task. *Anim. Behav.* **60,** 523–529 (2000).
35. Duguid, S. *et al.* Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proc. Biol. Sci.* **281,** 20141973 (2014).
36. Soc, P. R., St-Pierre, A., Larose, K. & Dubois, F. Prisoner's Dilemma Long-term social bonds promote cooperation in the iterated Subject collections Long-term social bonds promote cooperation in the iterated Prisoner's Dilemma. (2009). doi:10.1098/rspb.2009.1156
37. Massen, J. J. M., van den Berg, L. M., Spruijt, B. M. & Sterck, E. H. M. Generous leaders and selfish underdogs: pro-sociality in despotic macaques. *PLoS One* **5,** e9734 (2010).
38. Tomasello, M., Melis, A. P., Tennie, C., Wyman, E. & Herrmann, E. Two Key Steps in the Evolution of Human Cooperation. *Curr. Anthropol.* **53,** 673–692 (2012).
39. Okasha, S. On the Use of Rational Choice Theory in Evolutionary Biology. *pdfs.semanticscholar.org* at <https://pdfs.semanticscholar.org/a2d4/4ab4312e4fdf1eb52b00

c2c0264a01fdf888.pdf>

40. Raihani, N. & Bshary, R. Resolving the iterated prisoner's dilemma: theory and reality. *J. Evol. Biol.* **24,** 1628–1639 (2011).
41. De Martino, B., Kumaran, D., Seymour, B. & Dolan, R. J. Frames, biases, and rational decision-making in the human brain. *Science* **313,** 684–7 (2006).
42. Bargh, J. A., Schwader, K. L., Hailey, S. E., Dyer, R. L. & Boothby, E. J. Automaticity in social-cognitive processes. *Trends Cogn. Sci.* **16,** 593–605 (2012).
43. Fenici, M. Embodied Social Cognition and Embedded Theory of Mind. *Biolinguistics* **6,** 3–4 (2012).
44. Miner, B. G., Sultan, S. E., Morgan, S. G., Padilla, D. K. & Relyea, R. A. Ecological consequences of phenotypic plasticity. *Trends Ecol. Evol.* **20,** 685–692 (2005).
45. Lehner, S. R., Rutte, C. & Taborsky, M. Rats Benefit from Winner and Loser Effects. *Ethology* **117,** 949–960 (2011).
46. Eisenegger, C., Haushofer, J. & Fehr, E. The role of testosterone in social interaction. *Trends Cogn. Sci.* **15,** 263–271 (2011).
47. Creel, S. Social dominance and stress hormones. *Trends Ecol. Evol.* **16,** 491–497 (2001).
48. Møller, A. P. False Alarm Calls as a Means of Resource Usurpation in the Great Tit Parus major. *Ethology* **79,** 25–30 (2010).
49. Macy, M. W. Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange'. *Source*

Am. J. Sociol. **97,** 808–843 (1991).

50. Skyrms, B. The Stag Hunt. *Proc. Addresses Am. Philos. Assoc.* **75,** 31 (2001).
51. Sander Van Doorn, G., Riebli, T. & Taborsky, M. Coaction versus reciprocity in continuous-time models of cooperation. (2014). doi:10.1016/j.jtbi.2014.03.019
52. Alvard, M. S. & Nolin, D. A. Rousseau's Whale Hunt? *Curr. Anthropol.* **43,** 533–559 (2002).
53. Wood, R. I., Kim, J. Y. & Li, G. R. Cooperation in rats playing the iterated Prisoner's Dilemma game. *Anim. Behav.* **114,** 27–35 (2016).
54. Blanchard, R. J. & Caroline Blanchard, D. Aggressive behavior in the rat. *Behav. Biol.* **21,** 197–224 (1977).
55. Barnett, S. A. The rat: A study in behaviour. (1963).
56. Márquez, C., Rennie, S. M., Costa, D. F. & Moita, M. A. Prosocial choice in rats depends on food-seeking behavior displayed by recipients. *Curr. Biol.* **25,** 1736–1745 (2015).
57. Thomas R Zental & John M Levine. Observational Learning and Social Facilitation in the Rat. *Science (80-.).* **216,** 198–479 (1971).
58. Heyes, C. M. Social learning in animals: categories and mechanisms. *Biol. Rev. Camb. Philos. Soc.* **69,** 207–31 (1994).
59. Posadas-Andrews, A. & Roper, T. J. Social transmission of food-preferences in adult rats. *Anim. Behav.* **31,** 265–271 (1983).
60. Noble, J., Todd, P. M. & Tuci, E. Explaining social learning of food preferences without aversions: an evolutionary simulation model of Norway rats. *Proc. Biol. Sci.* **268,** 141–9 (2001).

61. Heyes, C. M., Ray, E. D., Mitchell, C. J. & Nokes, T. Stimulus Enhancement: Controls for Social Facilitation and Local Enhancement. *Learn. Motiv.* **31,** 83–98 (2000).

62. Pereira, A. G., Cruz, A., Lima, S. Q. & Moita, M. A. Silence resulting from the cessation of movement signals danger. *Curr. Biol.* **22,** R627–R628 (2012).

63. Masuda, A. & Aou, S. Social Transmission of Avoidance Behavior under Situational Change in Learned and Unlearned Rats. *PLoS One* **4,** e6794 (2009).

64. Galef, B. G. & Giraldeau, L.-A. Social influences on foraging in vertebrates: causal mechanisms and adaptive functions. *Anim. Behav. MS* **61,** 3–15 (2001).

65. Davis, H. Transitive inference in rats (Rattus norvegicus). *J. Comp. Psychol.* **106,** 342–349 (1992).

66. Uchida, N. & Mainen, Z. F. Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* **6,** 1224–1229 (2003).

67. Zeeb, F. D., Robbins, T. W. & Winstanley, C. A. Serotonergic and Dopaminergic Modulation of Gambling Behavior as Assessed Using a Novel Rat Gambling Task. *Neuropsychopharmacology* **34,** 2329–2343 (2009).

68. Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455,** 227–231 (2008).

69. Tervo, D. G. R. *et al.* Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell* **159,** 21–32 (2014).

70. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making.

Nat. Rev. Neurosci. **9,** 545–556 (2008).

71. Ben-Ami Bartal, I., Decety, J. & Mason, P. Empathy and pro-social behavior in rats. *Science* **334,** 1427–30 (2011).

72. Bartal, B.-A. *et al*. Pro-Social Behavior In Rats Requires An Affective Motivation. (2016). doi:10.1101/044180

73. Ben-Ami Bartal, I., Rodgers, D. A., Bernardez Sarria, M. S., Decety, J. & Mason, P. Pro-social behavior in rats is modulated by social experience. *Elife* **3,** e01385 (2014).

APPENDICES

Report

Current Biology

Prosocial Choice in Rats Depends on Food-Seeking Behavior Displayed by Recipients

Highlights

- We developed an automated two-choice task to study prosocial behavior in rats
- Rats showed a high proportion of prosocial choices in the absence of self-benefit
- We were able to manipulate our task to disentangle mechanisms of prosocial choice
- Recipients' food-seeking behavior is necessary to drive prosocial choice

Authors

Cristina Márquez, Scott M. Rennie, Diana F. Costa, Marta A. Moita

Correspondence

cristina.marquez@neuro.fchampalimaud.org (C.M.), marta.moita@neuro.fchampalimaud.org (M.A.M.)

In Brief

Márquez et al. show that rats behave prosocially in a food-foraging task. By manipulating the behavior of, and reward to, recipients of help, they disentangled the proximate mechanisms of prosocial choice, showing that rats are sensitive to food-seeking behavior and reward delivered to conspecifics.

Márquez et al., 2015, Current Biology *25*, 1736–1745
June 29, 2015
http://dx.doi.org/10.1016/j.cub.2015.05.018

CellPress

CellPress

Current Biology
Report

Prosocial Choice in Rats Depends on Food-Seeking Behavior Displayed by Recipients

Cristina Márquez,[1,*] Scott M. Rennie,[1] Diana F. Costa,[1] and Marta A. Moita[1,*]
[1]Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Avenida Brasilia, Lisbon 1400-038, Portugal
*Correspondence: cristina.marquez@neuro.fchampalimaud.org (C.M.), marta.moita@neuro.fchampalimaud.org (M.A.M.)
http://dx.doi.org/10.1016/j.cub.2015.05.018

SUMMARY

Animals often are prosocial, displaying behaviors that result in a benefit to one another [1–15] even in the absence of self-benefit [16–21] (but see [22–25]). Several factors have been proposed to modulate these behaviors, namely familiarity [6, 13, 18, 20] or display of seeking behavior [16, 21]. Rats have been recently shown to be prosocial under distress [17, 18] (but see [26–29]); however, what drives prosociality in these animals remains unclear. To address this issue, we developed a two-choice task in which prosocial behavior did not yield a benefit or a cost to the focal rat. We used a double T-maze in which only the focal rat controlled access to the food-baited arms of its own and the recipient rat's maze. In this task, the focal rat could choose between one side of the maze, which yielded food only to itself (selfish choice), and the opposite side, which yielded food to itself and the recipient rat (prosocial choice). Rats showed a high proportion of prosocial choices. By manipulating reward delivery to the recipient and its ability to display a preference for the baited arm, we found that the display of food-seeking behavior leading to reward was necessary to drive prosocial choices. In addition, we found that there was more social investigation between rats in selfish trials than in prosocial trials, which may have influenced the focals' choices. This study shows that rats provide access to food to others in the absence of added direct self-benefit, bringing new insights into the factors that drive prosociality.

RESULTS AND DISCUSSION

Prosocial Choice

It has been proposed that the neurobiological systems at the basis of social behaviors such as empathy and helping conspecifics may be shared across mammalian species [30]. Still, the scarcity of evidence for prosocial behavior in animal species amenable to mechanistic studies and the difficulty in establishing behavioral paradigms that allow the dissection of the proximal factors promoting or constraining prosocial behavior have hindered the search for the mechanisms of prosociality. Thus, the development of new paradigms to study social behaviors, such as prosocial choice, under highly controlled environments and using laboratory animal models, will greatly benefit the search for generic mechanisms underlying prosociality. Rats, a social species widely used in neuroscience, have been shown to cooperate in tasks testing for coordination and direct and generalized reciprocity [2, 8, 9, 11, 31], and to possess the cognitive capacity to engage in cooperation in the context of social dilemma games [32]. More recently, rats have been shown to release a conspecific from a restrainer, a prosocial act that may not involve a benefit to the focal [17, 18]. Nonetheless, what drives the release of the restrained rat remains unclear [26–29]. In addition, to our knowledge, there are no published attempts at probing this form of prosocial behavior in the absence of stress. Tasks that use food rewards allow for a better control over the behavior of subjects and the outcomes of particular action choices, permitting the disentangling of factors that drive prosocial behavior. Therefore, we set out to develop a reward-based task to study the mechanisms of prosocial behavior without self-benefit in Sprague-Dawley rats. The task we developed was inspired in classical two-choice tasks used to study prosociality in other species, such as a primates and corvids [19, 33, 34].

To this end, we developed a fully automated double T-maze in order to minimize interference by the experimenter while at the same time allowing for a precise control and detailed monitoring of the behavior of the interacting individuals. In each maze, a center arm gave access to two food-baited arms gated by automated doors. Pairs of non-food-deprived cage-mate rats were tested in our double T-maze (one per rat). For each pair, one rat was assigned to be the focal (the decision maker in our task) and the other the recipient (whose access to the rewarded arms depended on the focal). The focal could choose between the side that provided food only to itself (the focal received one food pellet and the recipient none; selfish choice) and the opposite side, which provided food to itself and the recipient rat (both focal and recipient received one food pellet; prosocial choice). Thus, prosocial choice did not imply an added benefit or a cost to the focal rat. Because in our task there was no role reversal (the focal never took the place of the recipient and vice versa), there was no room for reciprocal cooperation to emerge.

Since it has been shown in chimpanzees that behavioral displays of intention are required for the focal to provide help [16, 21], in our task recipient rats were trained to display food-seeking behavior. This corresponded to poking a nose port that controlled the automated doors giving access to the side arms of each maze. The nose ports were placed above each door, such that when a rat would poke the nose port, the door underneath would open (see Figures 1A–1C and S1).

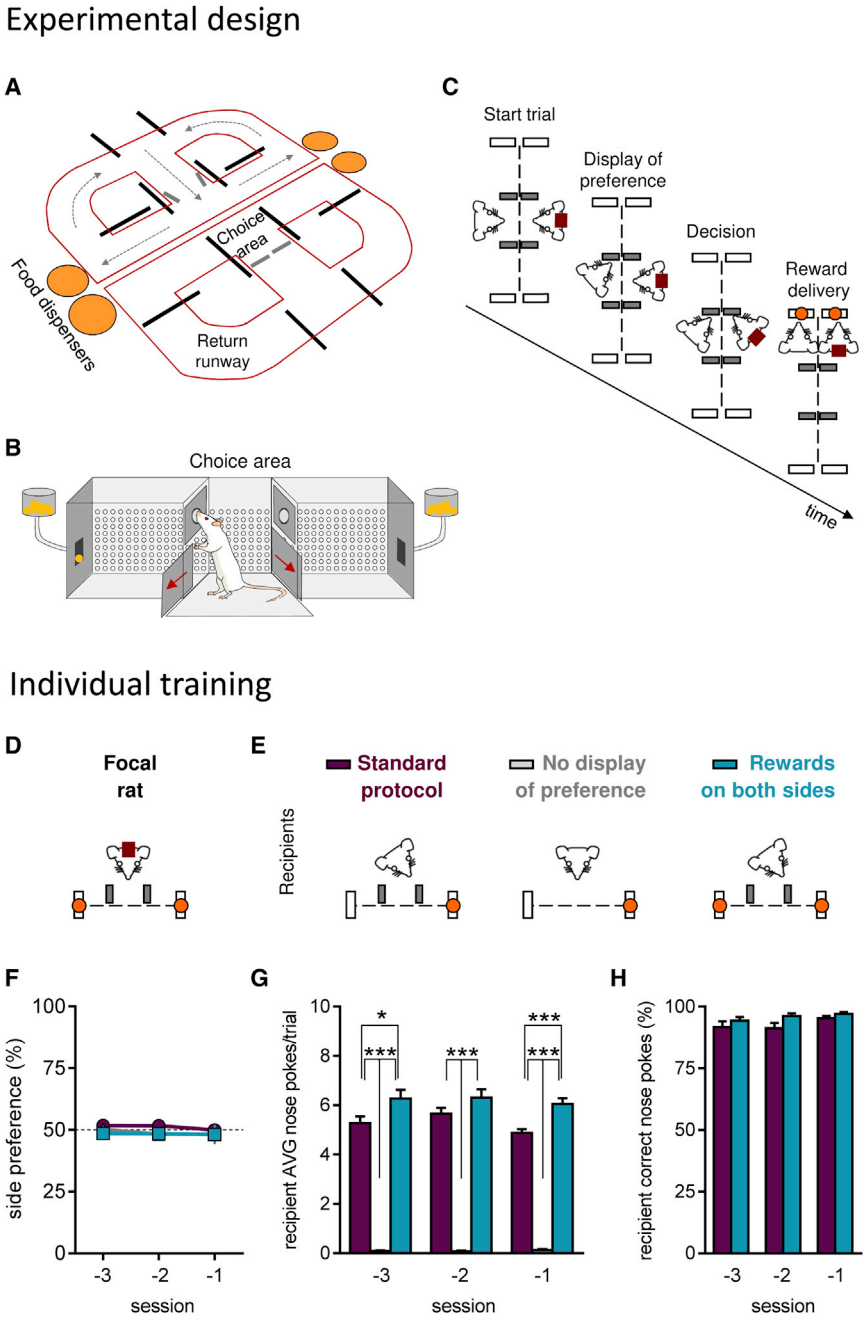


Figure 1. Two-Choice Task for the Study of Prosocial Behavior in Rats: Apparatus, Experimental Design, and Individual Training

The apparatus and experimental design of the prosocial choice task (A–C) and the individual training of rats prior to testing (D–H) are presented.

(A) Schematic view of the double T-maze. Each T-maze (one per rat), consisted of a center arm that gave access to two arms gated by automated doors (black lines), at the end of which food was delivered. Arrows in the upper maze represent the flow of movement of rats in the maze. Access to the choice area was prevented by automated doors placed in the central arm (gray lines).

(B) The doors in the choice area were controlled by nose ports placed above them, such that when a rat would poke the nose port the door underneath would open. See also Figure S1.

(C) Schematic view of a trial. A trial started when both rats were in the central arm. Once the recipient rat started displaying food-seeking behavior (nose pokes on preferred side), the nose ports of the focal rat became active and the focal rat could choose to nose poke on either side of its own maze. A single nose poke by the focal rat opened the door underneath the port and the door on that same side of the recipient's maze. Then, both animals entered the lateral arm and were rewarded according to experimental protocol. In the schema, the focal rat is represented with a red square, reward as orange circles, nose ports that control the opening of the doors as gray rectangles, and side preference as the angled head of the recipient toward the nose port of one side.

(D and E) The different individual training procedures for focal (D) and recipient (E) rats in "standard" and control protocols are schematized using the same symbols as in (C).

(F) Focal rats showed no side preference at the end of training, where side preference reports the percentage of choices to the arm that corresponded to the prosocial side during testing (since sides were counterbalanced across rats, for some animals side preference was the number of choices for the left side and for others the right side) (one-sample t test against chance for each experimental protocol independently, "standard" protocol: $t_{(14)} = 0.872$, $p' = 0.398$ for session −3; $t_{(14)} = 0.971$, $p' = 0.348$ for session −2; $t_{(14)} = -0.040$, $p' = 0.969$ for session −1; "no display of preference" protocol: $t_{(10)} = 0.105$, $p' = 0.918$ for session −3; $t_{(10)} = -1.082$, $p' = 0.304$ for session −2; $t_{(10)} = -1.037$, $p' = 0.324$ for session −1; "reward on both sides" protocol: $t_{(10)} = -0.763$, $p' = 0.463$ for session −3; $t_{(10)} = -0.713$, $p' = 0.491$ for session −2; $t_{(10)} = -0.471$, $p' = 0.647$ for session −1).

(G) The number of nose pokes per trial displayed by recipient rats differed across the three protocols. As expected, the number of nose pokes in the "no display of preference" condition was negligible. In addition, recipients of the "reward on both sides" protocol had a higher rate of nose pokes when compared to the "standard" condition (one-way ANOVA comparing experimental protocols in each training session, session −3: $F_{(2,34)} = 144.313$, $p < 0.00001$, further Fisher's least significant difference (LSD) post hoc tests revealed significant differences between "standard" and "reward on both sides" protocols [$p' = 0.011$] and between these two protocols and "no display of preference" protocol [$p' < 0.00001$]; session −2: $F_{(2,34)} = 192.864$, $p < 0.00001$, further LSD post hoc revealed a marginally significant difference between "standard" and "reward on both sides"' protocols [$p' = 0.058$] and a significant difference between these two protocols and "no display of preference" protocol [$p' < 0.00001$]; session −1: $F_{(2,34)} = 349.925$, $p < 0.00001$, further LSD post hoc revealed significant differences between "standard" and "reward on both sides" protocols [$p' < 0.0001$] and between these two protocols and "no display of preference" protocol [$p' < 0.00001$]).

(H) Recipients nose poked almost exclusively in the active port and in a similar manner in both conditions where animals were trained to display food-seeking behavior (independent sample t test for each training session: $t_{(24)} = 0.080$, $p = 0.937$ for session −3; $t_{(24)} = -1.057$, $p = 0.301$ for session −2; $t_{(24)} = -1.260$, $p = 0.220$ for session −1). Mean ± SEM are shown. *$p' < 0.05$, **$p' \leq 0.005$, ***$p' \leq 0.001$.

Corrected p values (p') after sequential Bonferroni correction for multiple comparisons are reported; three comparisons were performed for each protocol in (F) and each testing session in (G).

Before testing, both rats were trained individually to poke in the nose ports in order to open the doors of the food-baited arms, retrieve the food reward, and run around the maze to the choice area, initiating a new trial. Focals were trained until no side bias was observed. Training of recipients depended on experimental protocol (Figures 1D–1H and Supplemental Experimental Procedures). Once individual training was complete, rats were tested for prosocial behavior in the double T-maze for four daily test sessions of 40 min, during which rats completed an average of 32 trials.

During testing for prosocial behavior, both rats had access to the nose ports of their corresponding mazes; however, only the ports of the focal were active, and these controlled the doors of both mazes. In this manner, the recipient rat displayed food-seeking behavior (poking the deactivated port) while the focal controlled the recipient's access to the food-baited arms. It has previously been shown that under some circumstances rats seek proximity with conspecifics [24, 31, 35]. To avoid a contribution of the preference for being together to prosocial choice, in both prosocial and selfish choices, we ensured that focal and recipient rats always went to the same side of the maze, independently of whether recipients received reward or not. Importantly, rats could see, smell, hear, and partially touch each other through a transparent perforated wall that separated the two mazes. Hence, they could interact at the decision and reward areas. Finally, rats were synchronized such that the focal could always observe the recipient's food-seeking behavior at the choice point and its consumption of the food pellet at the reward area. To this end, the ports of the focal rat were only active once the recipient started poking its nose port, and the food pellet to the recipient would only be delivered once the focal rat had entered the reward area (for details, see Figure S1 and the Supplemental Experimental Procedures).

We first asked whether a rat (the focal) provides access to food to another rat (the recipient) in the absence of direct self-benefit. We found that rats quickly acquired a preference for the prosocial option, providing the recipient with access to the food-baited arm (Figure 2A, left panel, and Figure 2B, purple line). The proportion of prosocial choices was higher than chance in all testing sessions, already being significant on the first one (one-sample t test against chance, defined as 50% choice: $t_{(14)} = 3.14$, $p' = 0.014$ for session 1; $t_{(14)} = 2.64$, $p' = 0.019$ for session 2; $t_{(14)} = 4.91$, $p' = 0.001$ for session 3; $t_{(14)} = 4.31$, $p' = 0.002$ for session 4). This preference became higher and more reliable over the course of the daily test sessions (repeated-measures ANOVA with "test session" as within-subject factor: $F_{(3,42)} = 3.392$, $p = 0.027$; tests of within-subject contrasts revealed a linear effect: $F_{(1,14)} = 11.41$, $p = 0.005$). However, even though focals' choices at the end of individual training (baseline) were not different from chance (one-sample t test against chance: $t_{(14)} = 0.67$, $p' = 0.513$), it is still possible that small biases for a particular side of the maze could account for the observed preference for the prosocial side. To test this, we compared the proportion of prosocial choices during testing to the proportion of choices for the same side during individual training. A significant difference between baseline and the first prosocial test session was found, indicating that minor individual biases for a particular side of the maze does not account for the preference for the prosocial side observed during early testing (paired-sample t test, baseline against the first prosocial test session: $t_{(14)} = -2.90$, $p = 0.023$).

The preference for the prosocial side, apparent already in the first session, could result from a bias toward the prosocial side at the very beginning of testing possibly as a result of local enhancement or behavior copying. Thus, we further analyzed the dynamics of the preference for the prosocial choice by quantifying the proportion of prosocial choices for the first, middle, and last third of each session, focusing on the first session. We found that animals started at chance and a marginally significant bias for the prosocial side appeared toward the end of the first session (one-sample t test against chance: $t_{(14)} = 0.95$, $p' = 0.358$ for the first third of session 1; $t_{(14)} = 2.12$, $p' = 0.104$ for the middle third of session 1; $t_{(14)} = 2.66$, $p' = 0.056$ for the last third of session 1; see Figures 2C and S2). This result shows that focal rats rapidly, but gradually, acquired a preference for the prosocial side, possibly through learning of the contingency between their choice and the outcome to the recipient. Although most rats showed a reliable preference for the prosocial side (see Figure 2D, purple dots), the strength of this preference varied substantially ranging from 60% to 89% (for each rat, an exact test was performed in which the null hypothesis was that rats chose at chance, revealing that ten out of 15 had a significant preference for the prosocial side, one showed a preference for the selfish side, and four remained at chance).

Multiple factors could explain the observed preference for the prosocial choice that may or may not correspond to some form of other-regarding behavior. For example, reward delivery to the recipient could have triggered a reward signal in the focal's brain (vicarious reward), reinforcing prosocial choice. In addition, as observed in prior studies [16, 21], focals could be reacting to the display of food-seeking behavior of the recipients.

Sensitivity to the Display of Food-Seeking Behavior

To test the role of food-seeking behavior on prosocial choice we performed, in parallel, a second experiment with a different set of animals, in which recipient rats were not allowed to display a preference for the rewarded side before the focal made its choice. Recipient rats were held away from the choice area by a door in the center arm until the focal made its choice, being thus prevented from poking the nose ports or showing other forms of preference for the baited side (training rats to poke the nose port greatly decreased other behaviors at the choice area; still, sometimes we could observe behaviors such as gnawing on the door of the rewarded arm; see Movie S1). Therefore, in this experiment, the focal rat had to choose without the display of preference by the recipient rat. After the focal's decision, the central door that held the recipient rat opened, and the experiment proceeded in the same manner as in the "standard" condition. As in the "standard" protocol, recipient rats received food on only one side of the maze. Thus, access to food depended on the focal's choices ("no display of preference" condition; Figure 2A, middle panel). Figure 2B (gray line) shows that unlike in the first experiment, focal rats stayed at chance levels throughout all test sessions, showing no preference for the prosocial side (one-sample t test against chance: $t_{(10)} = -1.31$, $p' = 0.219$ for session 1; $t_{(10)} = -0.68$, $p' = 0.509$ for session 2; $t_{(10)} = -0.10$, $p' = 0.921$

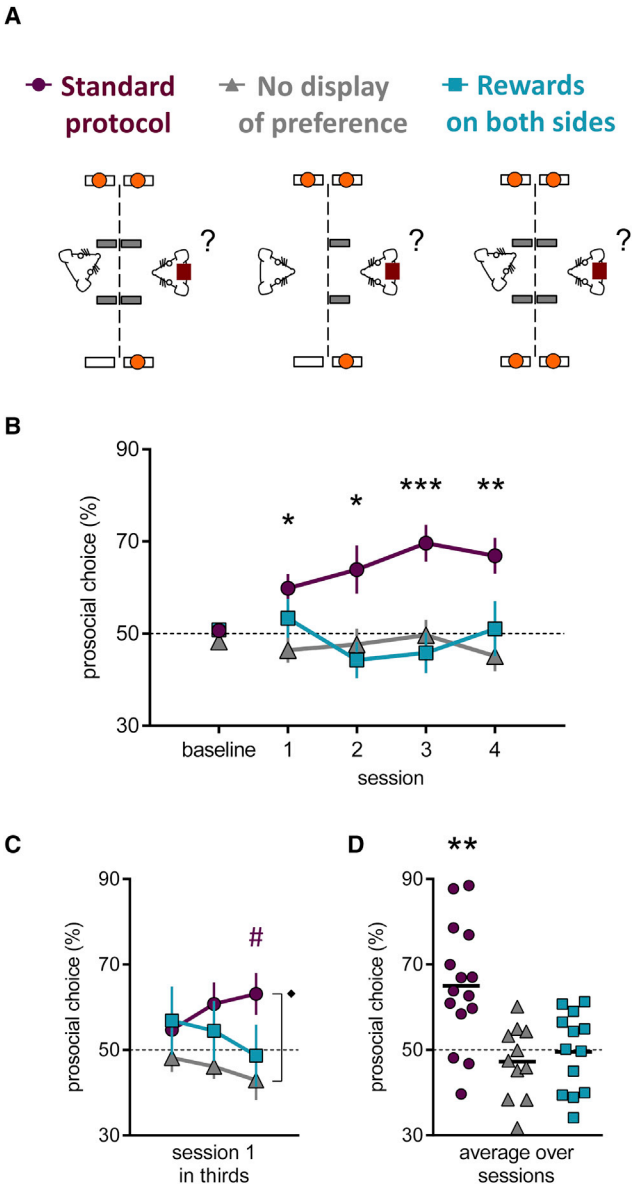


Figure 2. Prosocial Behavior in the Two-Choice Task

(A) Schematic views of the different protocols used to disentangle factors driving prosocial behavior are depicted using the same scheme as in Figure 1C. In all protocols, focal rats were always rewarded with one food-pellet, independently of choice, ensuring equal cost and benefit for both arms. In the "standard" protocol (left), recipients displayed food-seeking behavior by nose poking on the side where they would be rewarded. A prosocial choice (upper arm) resulted in reward to both animals, and a selfish choice (bottom arm) resulted in reward only to the focal animal. In the "no display of preference" protocol (middle), the recipient was prevented from accessing the choice area before the focal made its choice. Rewards were delivered as in the "standard" protocol—i.e., prosocial choice (upper arm), one food-pellet each, and selfish choice (bottom arm), one food-pellet to the focal rat only. Finally, in the "reward on both sides" protocol (right), the recipient still displayed food-seeking behavior toward one arm (in this case, the upper arm), but now one food-pellet was delivered to both animals on either side of the maze.

(B) Line graph shows that focal rats from the "standard" protocol (n = 15), but not the other two protocols (n = 11), quickly acquired a preference for the prosocial option, providing the recipient with access to the food-baited arm. Mean ± SEM of the percentage prosocial choices is shown for each experimental condition. Baseline and four test sessions are shown. Baseline corresponds to the percentage of choices for the arm that would later correspond to the prosocial side during testing, averaged across the last 2 days of individual training. One-sample t tests were used to compare proportion of prosocial choice in each test session within experimental groups against chance, and p values were adjusted after sequential Bonferroni correction. Repeated-measures ANOVA with "session" as a within-subject factor and "protocol" as a between-subject factor was performed to study differences between the different experimental protocols (see the main text).

(C) To study the emergence of prosocial choices within the first testing session, we divided performance in thirds. The percentage of prosocial choices for each third in each experimental protocol was then compared to chance (50%). No significant differences against chance were observed in the initial phase of testing, but a marginally significant preference toward prosocial choice emerged at the end of the session only in the "standard" protocol (see Figure S2 for further details). $^{\#}p' < 0.1$, deviation from chance; $^{\blacklozenge}p < 0.05$, difference between protocols.

(D) Individual values of prosocial choices averaged across sessions are plotted for each experimental condition. One-way ANOVA revealed differences between protocols, the "standard" protocol being significantly different from the other two conditions.

$^{*}p' < 0.05$, $^{**}p' \leq 0.005$, $^{***}p' \leq 0.001$. Corrected p values (p′) after sequential Bonferroni correction for multiple comparisons are reported; four comparisons we performed in (B) and three in (C). See also Figure S2 and Movie S1.

for session 3; $t_{(10)} = -1.46$, $p' = 0.176$ for session 4). This result shows that the recipient's food-seeking behavior was necessary to drive prosocial choice, which could be explained by a contribution of local enhancement, as observed in corvids [19], behavior copying, or a response to nose poking as a cue that may involve some form of mirror representation of the other's goals [21]. In addition, it shows that the cues related to reward delivery to the recipient rat, such as pellet odor, the sound of the pellet dropping, or chewing by the recipient, were not sufficient to drive prosocial choices. Finally, it suggest that although Norway rats in semi-natural conditions are tolerant to the presence of other rats at feeding sites [36, 37], feeding together does not by itself explain prosocial choices observed in the "standard" condition (first experiment). Analysis of deviation from chance for the choices of each focal rat showed that most rats remained at chance (seven out of 11), three displayed a preference for the selfish side, and only one rat showed a preference for the prosocial side. If eating together at least partially drove the focal rats' choices, a slight bias toward the prosocial side from some of the focal rats would be expected (a possibility even in the absence of a significant preference at the group level). This was, however, not the case.

Sensitivity to Reward Delivery

Our results show that focal rats are sensitive to the food-seeking behavior of the recipients, which could be the sole driver of prosocial choice. Alternatively, behavioral displays by recipients could be necessary, but not sufficient, to drive the preference for the prosocial side observed in the "standard" protocol. Focal rats may be sensitive to the reward delivered to the recipient rat, which together with the food-seeking behavior could drive prosocial choices. Indeed, it has been recently shown that upon the observation of a conspecific eating food, there is an initial increase in dopamine in the nucleus accumbens [38]. To test whether prosocial choices were driven solely by recipients'

food-seeking behavior and whether local enhancement or behavior copying could account for the focals' decisions, we ran a third experiment in another set of animals in which recipient rats still poked the nose port on one side only but received food pellets on both sides of the maze (see Figure 2A, right panel). In this manner, food-seeking behavior (poking the nose port on one side) was dissociated from the delivery of food (on both sides). If behavior copying or local enhancement explains the preference for the prosocial side observed in the "standard" protocol, then in this experiment focal rats should also show a preference for the side at which recipient rats poked. In contrast, if food delivery to the recipient rat also contributed to prosocial choice, in this experiment focal rats should remain at chance. In order to dissociate reward delivery from the side of nose poking, we changed the last 3 days of individual training of recipient rats, ensuring that they learned that poking the port on one side only was necessary to receive reward on either side of the maze (see Figure 1 and the Supplemental Experimental Procedures). In this third experiment, the behavior of the focal rat always led to reward to the recipient. Still, for simplicity and comparison with the previous experiments, we called the side of recipient nose poking the prosocial side. We found that focal rats stayed at chance level (Figure 2B, blue line), showing no preference for the side where the recipient was poking, conventionally called the prosocial side (one-sample t test against chance: $t_{(10)} = 0.79$, $p' = 0.449$ for session 1; $t_{(10)} = -1.44$, $p' = 0.181$ for session 2; $t_{(10)} = -0.94$, $p' = 0.368$ for session 3; $t_{(10)} = 0.17$, $p' = 0.871$ for session 4). This result suggests that rats were sensitive to the reward delivered to their cage-mates and that behavioral copying and local enhancement were not sufficient to sustain prosocial choices.

Comparison across Experimental Conditions

Next, we compared directly the choices of focal rats across the three experimental protocols over the course of the four testing days. To this end, we performed a repeated-measures ANOVA with "protocol" as a between-subjects factor and "test session" as a within-subjects factor. This analysis revealed a significant overall effect of "protocol" ($F_{(2,34)} = 10.01$, $p = 0.0004$), no effect of "test session" ($F_{(3,102)} = 0.623$, $p = 0.602$) and a marginally significant interaction between "protocol" and "test session" ($F_{(6,102)} = 1.945$, $p = 0.081$). Post hoc analysis revealed that rats in the "standard" protocol showed significantly higher levels of prosocial choices as compared to the two other behavioral protocols, except for session 1, where it differed from the "no display of preference" ($p' = 0.023$), but not the "reward on both sides" protocols ($p' = 0.179$) (which in turn was not significantly different from the "no display of preference" protocol [$p' = 0.183$]; see Figure 2B). These results further support the finding that food-seeking behavior is necessary, but not sufficient, to drive prosocial choices in focal rats and that focal animals are sensitive to the reward received by recipients. In addition, in the first session, rats in the "no display of preference" group showed lower levels of prosocial choices than rats in the other two groups (although the difference relative to rats in the "reward on both sides" protocol did not reach significance), suggesting that at early stages of testing local enhancement or behavior copying (triggered by the food-seeking behavior of recipients) may have facilitated learning. This would, however, not be sufficient to sustain a preference for the prosocial side, as only rats in the "standard" protocol reliably and gradually acquired this preference.

We also asked whether prosocial choice was a goal directed or a habitual choice. To this end, we trained an independent set of animals in a similar manner as for the "standard" protocol, and, after 4 days of testing, we changed contingencies by providing the recipient with reward on both sides (instead of reward delivery only on the prosocial side). Importantly, reward to the focal rat remained unchanged. If prosocial choice were a habitual action, we would expect focal rats to keep their preference for the prosocial side despite the change in the contingency between the focals' choice and reward delivered to recipients. However, we observed a drop in the preference for the prosocial choice when recipients were suddenly rewarded on both sides (see Figure S2B).

The apparent sensitivity of focal rats to the delivery of reward to recipients, as seen by the lack of preference for the prosocial side when rewards to recipients were delivered on both sides of the maze, may have resulted from the fact that recipient rats could be less motivated to display food-seeking behavior as they would be rewarded on both arms and, despite their training, poked less on the "prosocial" side to which they were trained. To address this issue, we quantified the number of pokes on the prosocial side displayed by recipient rats before the focal made its decision in the first ("standard" condition) and third ("reward on both sides") experiments, and we found no difference (Figure 3A; Mann-Whitney U test, $U = 57.5$, $p' = 0.198$ for average prosocial pokes over sessions). Pokes on the opposite, "selfish," side were negligible in both experiments. Moreover, the number of pokes made by recipients was similar between prosocial and selfish trials (data not shown). Hence, recipients' food-seeking behavior, expressed as nose poking on the trained side, was similar across the two experimental conditions and across trial type. This is expected since in both protocols during individual training, recipient rats had to poke on average five times in order to gain access to food, and during testing nose poking by the recipient was necessary to activate the focal's nose ports and thus to progress within the trial. We also quantified the number of times the recipient rat investigated the focal rat during the display of food-seeking behavior (between the start of the trial and the focal's decision) and again found no difference between the two protocols (see Social Interactions below and Figure S4). This result further supports the similarity in the recipients' behavior before the decision was made, to the extent that we could quantify, across the two protocols.

Sensitivity to Reward Rate and Temporal Discounting

We noticed that in some trials of the "standard" condition, where the focal chose the selfish choice, recipients were reluctant to enter the arm, thus delaying the beginning of the next trial (which required both rats to be back in the center arm). Indeed, the time elapsed between the moment the focal opened the doors and the entry of both rats in the reward area was longer for selfish than prosocial trials in the "standard" and "no display of preference" protocols (Figure S3A). These correspond to the two experimental conditions in which reward was delivered to the recipient on one side only. The systematic delay in selfish trials could lead to fewer trials, and therefore fewer rewards, on those

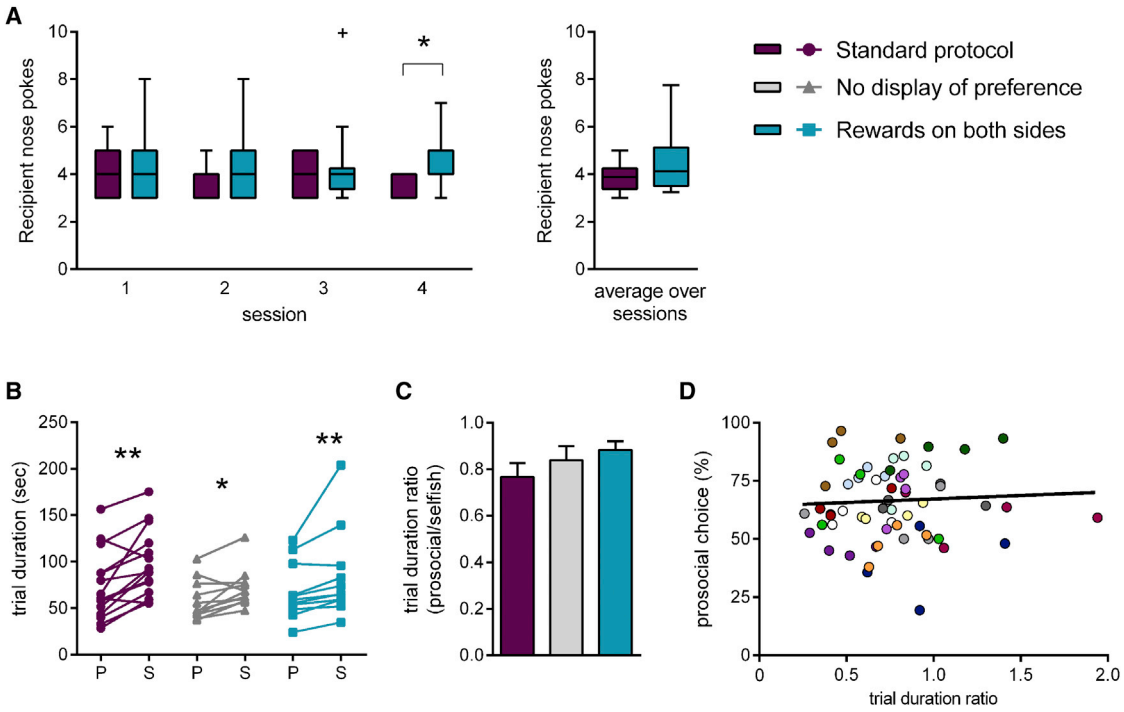


Figure 3. Recipients Display Similar Pre-decision Food-Seeking Behavior across Protocols and Focal's Reward Rate Does Not Predict Choice

(A) The number of nose pokes in the rewarded side displayed by recipients, before the focal made its choice, was similar in the "standard" and "reward on both sides" protocols during each test session (left) and averaged across sessions (right), except for the last session of prosocial testing, in which recipients of the "standard" protocol nose poked significantly less (Mann-Whitney test: $U = 81.5$, $p' = 0.959$ for session 1; $U = 59.5$, $p' = 0.237$ for session 2; $U = 74.5$, $p' = 0.683$ for session 3; $U = 31.0$, $p' = 0.024$ for session 4; $U = 57.5$, $p = 0.198$ for recipient pokes averaged over sessions).

(B) The median duration of selfish trials (S) was higher than that of prosocial trials (P) in all three protocols (paired-sample Wilcoxon signed-rank test within each experimental protocol: $Z = -3.010$, $p' = 0.008$ for the "standard" protocol; $Z = -2.311$, $p' = 0.021$ for the "no display of preference" protocol; $Z = -2.845$, $p' = 0.009$ for the "reward on both sides" protocol). See also Figure S3.

(C) When the ratio between the duration of prosocial and selfish trials was calculated for each individual animal, no differences between protocols were observed (one-way ANOVA: $F_{(2,34)} = 1.154$, $p = 0.327$).

(D) Scatter plot showing percentage of prosocial choice and prosocial/selfish trial duration ratio (for each rat, four data points are shown, one per session). Pearson's correlation between the two variables was not significant, confirming that trial duration (or reward rate) was not predictive of choice. Different animals are color coded.

In (A), box plots show median, first and third quartiles, and minimum and maximum values. Crosses represent outliers. Mean ± SEM are shown in (C). *$p' < 0.05$, **$p' \leq 0.005$. Corrected p values (p') after sequential Bonferroni correction for multiple comparisons are reported; four comparisons were performed in (A) and three in (B).

sessions where more "selfish" choices were made, raising the possibility that animals in the "standard" protocol chose the prosocial side more often in order to maximize their own reward [39]. We compared the number of trials across protocols for all sessions and found no differences in the total number of trials performed by rats in the three experimental protocols; hence, focal rats in the different protocols received a similar number of rewards (Figure S3B). Still, longer selfish trials imply (1) a higher reward rate for rats choosing more often the prosocial side and (2) a delay (relative to prosocial trials) in starting a new trial and hence the opportunity to get another reward. This could lead to temporal discounting (i.e., attribution of lower value to delayed reward) of food reward on the selfish side, which in turn could have influenced the choices of focal rats. Indeed, temporal discounting has been proposed as a crucial factor in social decision-making and cooperative behavior [40].

Therefore, we compared median trial duration of prosocial and selfish choice trials in all three conditions. Selfish trials were significantly longer than prosocial trials in all three experimental protocols (Wilcoxon signed-rank test within each protocol: $Z = -3.01$, $p' = 0.008$ for the "standard" protocol; $Z = -2.31$, $p' = 0.021$ for the "no display of preference"; and $Z = -2.85$, $p' = 0.009$ for the "reward on both sides" conditions) (Figure 3B). As variability in trial duration across interacting dyads could mask differences between experimental conditions, we computed the ratio of median trial duration between prosocial and selfish trials for each rat. Still, no difference was found across conditions (Figure 3C; one-way ANOVA: $F_{(2,34)} = 1.154$, $p = 0.327$). This finding indicates that the observed increase in reward rate on the prosocial side was not sufficient to drive a preference for that side, as in both the "no display of preference" and "reward on both sides" conditions no preference for the prosocial side was seen despite the increased reward rate associated with choosing that side. It was still possible that the difference observed between prosocial and selfish trial duration could explain the variance observed in the proportion of prosocial

choices displayed by focals in the ''standard'' condition (see Figure 2D). To examine this possibility, we performed a Pearson's correlation across the proportion of prosocial choice (displayed by each rat on each of the four test sessions) and the ratio between prosocial and selfish trial duration. We found no correlation between these two measures ($r = 0.030$, $p = 0.82$; Figure 3D). In addition, the relative difference in trial duration only emerged in session 3 (Figures S3C–S3F), whereas a preference for the prosocial side was apparent in the very first test session (see Figure 2B).

However, since rodents have been shown to perceive absolute time differences [41], it is possible that in our task the focal rats were not sensitive to the relative difference between prosocial and selfish trial duration (as calculated with the ratio), but rather were responding to absolute time differences which could differ between experimental protocols. To evaluate this possibility, we calculated the median difference in trial duration between trial types focusing on the two experimental protocols in which recipients were rewarded only in one side (i.e., ''standard'' and ''no display of preference'' protocols). We found that (1) absolute time difference in trial duration was not significantly different between experimental protocols over days (a repeated-measures ANOVA with ''experimental protocol'' as a between-subject factor and ''session'' as a within-subject factor revealed a significant effect of ''session'' [$F_{(3,78)} = 5.959$, $p = 0.0.001$], no significant effect of ''experimental protocol'' [$F_{(1,26)} = 1.205$, $p = 0.282$], and no significant interaction of testing session by experimental protocol [$F_{(3,78)} = 0.206$, $p = 0.892$]), (2) as observed with the ratio, absolute time differences emerged later than a preference for the prosocial choice, being only significant from the third testing session onward (data not shown), and that (3) there was no correlation between choice and absolute time difference ($r = -0.050$, $p = 0.703$ for the ''standard'' protocol; $r = -0.175$, $p = 0.225$ for the ''no display of preference'' protocol).

Taken as a whole, these findings make reward rate and temporal discounting as the main driving force of prosocial choice highly unlikely. The fact that time difference (absolute or relative) between prosocial and selfish trials emerged later than the preference for prosocial choice indicates that reward rate and temporal discounting are not sufficient to initiate prosocial choice. This is possibly due to the fact that rats were not food deprived and that in all experiments rewards were available to the focal rat as soon as it made its choice, such that the delay to receive the reward (from nose poking to pellet retrieval from the food magazine) was independent of the recipient's behavior and hence of trial type. However, we cannot exclude the possibility that these factors could play a role in the maintenance of choice in later testing sessions.

Social Interactions

In all experiments, rats had ample opportunity to interact, which could affect the decisions made by focal rats. Therefore, we examined how the interactions between focal and recipient rats might relate to the focals' choices. To this end, we quantified the number of times focal and recipient rats investigated each other, the number of times the focal unilaterally investigated the recipient, and vice versa. Given that the three forms of social investigation followed the same pattern (see Figures S4A–S4D for an analysis of each type of interaction), we used the sum of all three forms of interaction, henceforth referred to as social investigation. The amount of social investigation was similar in the ''standard'' and ''reward on both sides'' protocol, being higher than in the ''no display of preference'' protocol, as in the latter group rats could not interact before the focal made its decision (Figure 4A, left panel). In addition, for all three conditions, the number of interactions was highest on the first test session (Figure 4A, right panel).

One possibility is that social investigation is not homogenous across different stages of the trial, such as before and after the focal's decision, and differences across experimental protocol would only be apparent in a particular segment of the trial. Hence, we next divided the trials in three segments and analyzed the number of social investigation bouts in each of the following: (1) investigation bouts taking place before the focal made its choice (only possible in the ''standard'' and the ''reward on both sides'' conditions), (2) investigation bouts from the moment the focal made its decision until the recipient retrieved its food (or in selfish trials until the recipient entered the lateral unrewarded arm), and (3) investigation bouts from the moment the recipient retrieved its reward (or in selfish trials until the recipient entered the lateral unrewarded arm) until both rats were back to the central arm to initiate another trial. We found that the number of social investigation bouts was similar across protocols in the different trial segments, except for the increased number of social investigation in rats of the ''no display of preference'' condition during the time between the focals' decision and the recipients' retrieval of food. This is possibly because rats in this experimental group could not interact before the focal made its decision (one-way ANOVA: $F_{(1,24)} = 1.447$, $p = 0.241$ for ''prior to decision''; $F_{(2,33)} = 19.477$, $p < 0.0001$ for ''from decision to recipient's reward''; LSD post hoc: ''standard'' against ''no display of preference,'', $p' < 0.0001$; ''standard'' against ''reward on both sides,'', $p' = 0.068$; ''reward on both sides'' against ''no display of preference,'' $p' < 0.0001$; $F_{(2,33)} = 1.780$, $p = 0.184$ for ''from recipient's reward to trial end''; Figure 4B).

Finally, we assessed differences in social investigation between prosocial and selfish trials for all experimental protocols, as these could contribute to the decisions of focal rats. We found that prior to the focal rats' decision, the number of social investigation bouts was similar across trial type (Figure S4E), suggesting that social interactions before the focal rat made its decision were not related to prosocial choice. However, after the focal rat made its decision, rat dyads in the ''standard'' and ''no display of preference'' protocols displayed higher levels of social investigation in selfish trials relative to that observed in prosocial trials (paired-sample Wilcoxon signed-rank tests between prosocial and selfish trials for each independent protocol: $Z = -3.408$, $p' = 0.001$ for the ''standard'' protocol; $Z = -2.803$, $p' = 0.015$ for the ''no display of preference'' protocol; $Z = -1.156$, $p' = 0.248$ for the ''reward on both sides'' protocol; Figure 4C, left panel). We calculated the ratio between the amount of social investigation displayed in prosocial and selfish trials for each protocol and found that this ratio was highest in the ''standard'' protocol (one-way ANOVA: $F_{(2,33)} = 31.478$, $p < 0.0001$; LSD post hoc: ''standard'' against ''no display of preference'' protocol, $p' = 0.004$;

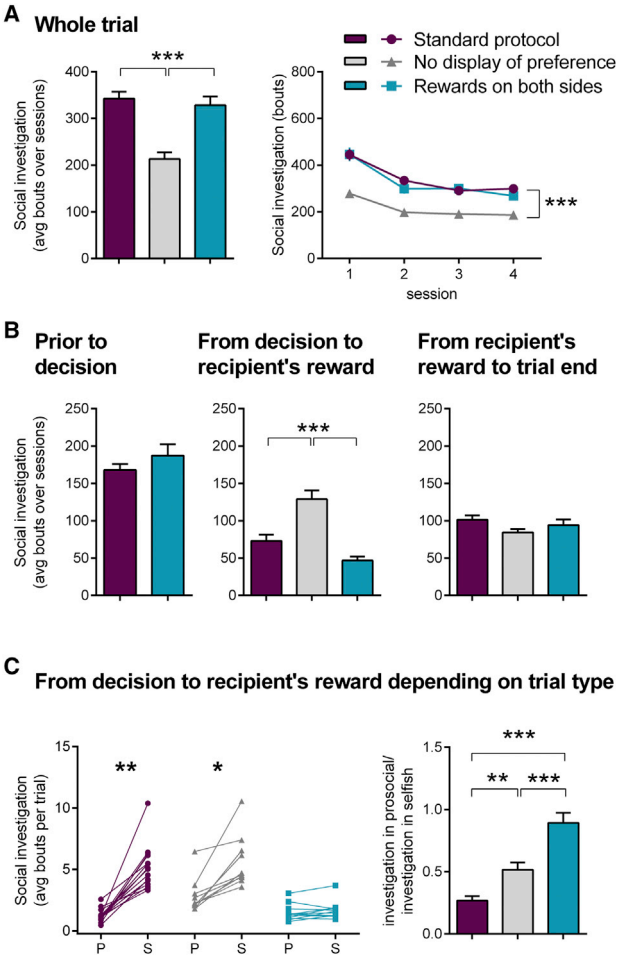


Figure 4. Social Investigation after the Focal Makes Its Decision Is Higher in Selfish Than Prosocial Trials

(A) The amount of social investigation between animals was lower in the "no display of preference" protocol compared to the "standard" and "reward on both sides" protocols, which did not differ between each other. The number of social investigation bouts decreased over sessions in a similar manner in all experimental protocols (repeated-measures ANOVA with "session" as a within-subject factor and "protocol" as a between-subjects factor confirmed differences between the protocols [$F_{(2,33)} = 17.73$, $p < 0.00001$]; further LSD post hoc confirmed lower levels of social interaction in all testing sessions compared to the "standard" [$p' = 0.00001$] and "reward on both sides" [$p' = 0.00009$] protocols) and a decrease of the amount of social investigation bouts over testing sessions ($F_{(3,99)} = 59.184$, $p < 0.00001$) that was similar in all experimental protocols ("session" × "protocol" $F_{(6,99)} = 2.036$, $p = 0.068$).

(B) No differences in the amount of social investigation (1) prior to the focal's decision, (2) from the moment of the decision until both animals were in the reward areas, and (3) from reward to trial end were observed between the protocols, except for the increased number of social investigation in rats of the "no display of preference" protocol in (2).

(C) Rat dyads from the "standard" and "no display of preference" protocols showed increased social investigations in selfish trials from the moment of decision (the focal pokes the nose port) until the recipient's reward delivery. This difference in the number of social investigation bouts was more pronounced in the "standard" protocol, as seen by the ratio between social investigation in prosocial and selfish trials. See also Figure S4.

Mean ± SEM are shown in all panels, except for (C), where individual values are plotted. *$p' < 0.05$, **$p' \leq 0.005$, ***$p' \leq 0.001$. Corrected p values (p') after sequential Bonferroni correction for multiple comparisons are reported; three comparisons were performed in each graph.

"standard" against "reward on both sides" protocol, $p' < 0.0001$, "reward on both sides" against "no display of preference" protocol, $p' = 0.0002$; Figure 4C, right panel). Since the difference in social investigation was observed in the two protocols in which the recipient received food only on one side, this result suggests that reward delivery influenced the way in which rats interact, which in turn may have influenced prosocial behavior. More specifically, the lack of reward delivery to recipient rats in selfish trials may have driven more social investigation by the dyad. The fact that a similar pattern is observed for all three forms of social investigation—mutual investigation, focal investigates recipient, and vice versa—suggests that when one rat investigates the other, the target of investigation tends to investigate back (see Figure S4G). Given the limitations in the accuracy of our behavioral analysis for the short investigation bouts that we measured (average bout duration: 0.97 ± 0.017 s), we could not determine who was driving the interaction, the focal or the recipient rat. Still, as it is the difference in reward received by the recipient that explains the differences in social investigation, it is possible that the observed increase in social investigation in selfish trials at the reward area may have been driven by recipient rats, which could be akin to begging or harassment displays observed in primates [1, 4, 22]. These interactions, together with the display of food-seeking behavior (nose poking), may have led to increased prosocial choices by focal rats in the "standard" condition.

Conclusions

We found that rats were prosocial in a food-foraging task, providing access to food to a cage-mate in the absence of added self-benefit or cost. This behavior was modulated by the display of food-seeking behavior expressed as poking a nose port on the door that gave access to the food-baited arm and by social interactions during the task.

By demonstrating that rats provide food to others in the absence of a direct or deferred benefit within the context of the task, we expand previous studies on prosocial behavior in rats reporting that these animals provide food to others in reciprocity-based tasks [8, 9] and that they relieve others from stress in the absence of self-benefit [17, 18]. Attempts to find evidence of provision of food to others under laboratory settings have provided conflicting evidence, possibly due to the fact that for animals to display this form of prosocial behavior they may have to overcome the drive to compete for food (even when subjects are not competing for food directly) [16, 42]. Rats have been shown to compete for food in a foraging task [43]; however, early descriptions of rats' natural behavior report that these animals are often found feeding in groups, showing tolerance for the presence of others at a food site even in conditions of limited food resources [36, 37].

Consistent with previous studies in primates [16, 21], we found the display of food-seeking behavior—in our task, poking in the nose port that gave access to the food-baited arm—to be crucial for prosocial choices by focal rats. Focal animals followed the recipient's nose-poking behavior to make their choices, this factor being necessary for the emergence of prosocial choice. To our knowledge, this is the first evidence of gaze (body)-following behavior in rats, which could be an

important component of social coordination. Moreover, prosocial behavior was modulated by reward delivery to the recipient rat. Therefore, in our task, displays of food-seeking behavior were not sufficient to drive prosocial behavior. In addition, we found that focal rats were sensitive to a change in the contingency between their action and reward to the recipient rats, suggesting that the focals' choices were goal directed. The specific mechanism and sensory cues by which displays of food-seeking behavior and sensitivity to reward to a conspecific modulate prosocial choice remain to be established. Our results raise the possibility that vicarious reward signals reinforced prosocial choice. Vicarious reward signals have been shown in the brain of human and non-human primates [44–46]. Furthermore, recent reports show that the observation of a conspecific eating food drives an initial dopamine increase in the nucleus accumbens of rats [38] and that social reward signals are mediated by oxytocin and serotonin in the nucleus accumbens of mice [47]. Interestingly, vicarious reward signals are modulated by oxytocin in monkeys [48]. In addition to inducing vicarious reward signals, food delivery may have driven prosocial choice by altering the behavior of recipient rats. Indeed, this was the case in our task, since in selfish trials rats took longer to enter the reward area and displayed more bouts of social investigation. Whether vicarious reward signals in the rat brain can drive prosocial behavior and how the behavior of the recipient leads to prosocial choice remain to be established.

We believe that using a classical type of decision-making task in combination with the vast tools available in rodents to record and manipulate brain activity will greatly impact the search of the neural mechanism underlying prosocial behavior.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, two tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2015.05.018.

AUTHOR CONTRIBUTIONS

C.M. and M.A.M. conceived and designed the changes made to the automated double T-maze previously developed by S.M.R. to study cooperative behavior using two-choice tasks in rats. C.M. and M.A.M. designed all experiments. C.M. established the behavioral protocols. C.M. and D.F.C. performed all experiments. D.F.C. performed the offline video scoring of social interactions. S.M.R. analyzed the data regarding nose pokes, parsed the data obtained from the automated mazes, and aligned it to the video data. C.M. performed all statistical analysis. All authors discussed the data, and M.A.M. and C.M. wrote the paper.

ACKNOWLEDGMENTS

We would like to thank Eric DeWitt for his technical assistance in programming for data analysis, the Scientific Software Platform for the code to compute trial duration, Tomás Cruz for creating the custom-made program for offline scoring of video recorded behavior, the Scientific Hardware platform for excellent technical assistance on the development of the mazes, the personnel of the animal facility for the care of our experimental animals, and Wieland Brendel for the photographs of the rats in the maze. We would also like to acknowledge the Behavioral Neuroscience Laboratory, Léa Zink, Claudia Feierstein, Susana Lima, Maria Luisa Vasconcelos, and Alfonso Renart for fruitful discussions of the project and Joseph Paton for comments on the manuscript. This work was funded by the Champalimaud Foundation.

Received: June 19, 2014
Revised: April 20, 2015
Accepted: May 11, 2015
Published: June 4, 2015

REFERENCES

1. Burkart, J.M., Fehr, E., Efferson, C., and van Schaik, C.P. (2007). Other-regarding preferences in a non-human primate: common marmosets provision food altruistically. Proc. Natl. Acad. Sci. USA *104*, 19762–19766.
2. Daniel, W.J. (1942). Cooperative problem solving in rats. J. Comp. Psychol. *34*, 361–368.
3. Drea, C., and Carter, A. (2009). Cooperative problem solving in a social carnivore. Anim. Behav. *78*, 967–977.
4. Horner, V., Carter, J.D., Suchak, M., and de Waal, F.B. (2011). Spontaneous prosocial choice by chimpanzees. Proc. Natl. Acad. Sci. USA *108*, 13847–13851.
5. Massen, J.J., van den Berg, L.M., Spruijt, B.M., and Sterck, E.H. (2010). Generous leaders and selfish underdogs: pro-sociality in despotic macaques. PLoS ONE *5*, e9734.
6. Owens, D.D., and Owens, M.J. (1984). Helping behaviour in brown hyenas. Nature *308*, 843–845.
7. Plotnik, J.M., Lair, R., Suphachoksahakun, W., and de Waal, F.B. (2011). Elephants know when they need a helping trunk in a cooperative task. Proc. Natl. Acad. Sci. USA *108*, 5116–5121.
8. Rutte, C., and Taborsky, M. (2007). Generalized reciprocity in rats. PLoS Biol. *5*, e196.
9. Rutte, C., and Taborsky, M. (2008). The influence of social experience on cooperative behavior of rats (Rattus norvegicus): direct vs generalised reciprocity. Behav. Ecol. Sociobiol. *62*, 499–505.
10. Scheid, C., and Noë, R. (2010). The performance of rooks in a cooperative task depends on their temperament. Anim. Cogn. *13*, 545–553.
11. Schneeberger, K., Dietz, M., and Taborsky, M. (2012). Reciprocal cooperation between unrelated rats depends on cost to donor and benefit to recipient. BMC Evol. Biol. *12*, 41.
12. Seed, A.M., Clayton, N.S., and Emery, N.J. (2008). Cooperative problem solving in rooks (Corvus frugilegus). Proc. Biol. Sci. *275*, 1421–1429.
13. St-Pierre, A., Larose, K., and Dubois, F. (2009). Long-term social bonds promote cooperation in the iterated Prisoner's Dilemma. Proc. Biol. Sci. *276*, 4223–4228.
14. Suchak, M., and de Waal, F.B. (2012). Monkeys benefit from reciprocity without the cognitive burden. Proc. Natl. Acad. Sci. USA *109*, 15191–15196.
15. Taborsky, M. (2013). Social evolution: reciprocity there is. Curr. Biol. *23*, R486–R488.
16. Yamamoto, S., Humle, T., and Tanaka, M. (2009). Chimpanzees help each other upon request. PLoS ONE *4*, e7416.
17. Ben-Ami Bartal, I., Decety, J., and Mason, P. (2011). Empathy and pro-social behavior in rats. Science *334*, 1427–1430.
18. Ben-Ami Bartal, I., Rodgers, D.A., Bernardez Sarria, M.S., Decety, J., and Mason, P. (2014). Pro-social behavior in rats is modulated by social experience. eLife *3*, e01385.
19. Schwab, C., Swoboda, R., Kotrschal, K., and Bugnyar, T. (2012). Recipients affect prosocial and altruistic choices in jackdaws, Corvus monedula. PLoS ONE *7*, e34922.
20. Tan, J., and Hare, B. (2013). Bonobos share with strangers. PLoS ONE *8*, e51922.
21. Warneken, F., Hare, B., Melis, A.P., Hanus, D., and Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. PLoS Biol. *5*, e184.
22. Brosnan, S.F., and De Waal, F.B. (2003). Monkeys reject unequal pay. Nature *425*, 297–299.

23. Di Lascio, F., Nyffeler, F., Bshary, R., and Bugnyar, T. (2013). Ravens (Corvus corax) are indifferent to the gains of conspecific recipients or human partners in experimental tasks. Anim. Cogn. *16*, 35–43.
24. Silberberg, A., Allouch, C., Sandfort, S., Kearns, D., Karpel, H., and Slotnick, B. (2014). Desire for social contact, not empathy, may explain "rescue" behavior in rats. Anim. Cogn. *17*, 609–618.
25. Skerry, A.E., Sheskin, M., and Santos, L.R. (2011). Capuchin monkeys are not prosocial in an instrumental helping task. Anim. Cogn. *14*, 647–654.
26. Balter, M. (2012). Animal cognition. 'Killjoys' challenge claims of clever animals. Science *335*, 1036–1037.
27. Heyes, C. (2012). Simple minds: a qualified defence of associative learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 2695–2703.
28. Hollis, K.L., and Nowbahari, E. (2013). A comparative analysis of precision rescue behaviour in sand-dwelling ants. Anim. Behav. *85*, 537–544.
29. Vasconcelos, M., Hollis, K., Nowbahari, E., and Kacelnik, A. (2012). Prosociality without empathy. Biol. Lett. *8*, 910–912.
30. Decety, J. (2011). The neuroevolution of empathy. Ann. N Y Acad. Sci. *1231*, 35–45.
31. Schuster, R., and Perelberg, A. (2004). Why cooperate? An economic perspective is not enough. Behav. Processes *66*, 261–277.
32. Viana, D.S., Gordo, I., Sucena, E., and Moita, M.A. (2010). Cognitive and motivational requirements for the emergence of cooperation in a rat social game. PLoS ONE *5*, e8483.
33. Colman, A.D., Liebold, K.E., and Boren, J.J. (1969). A method for studying altruism in monkeys. Psychol. Rec. *19*, 401–405.
34. Silk, J.B., Brosnan, S.F., Vonk, J., Henrich, J., Povinelli, D.J., Richardson, A.S., Lambeth, S.P., Mascaro, J., and Schapiro, S.J. (2005). Chimpanzees are indifferent to the welfare of unrelated group members. Nature *437*, 1357–1359.
35. Tsoory, M.M., Youdim, M.B., and Schuster, R. (2012). Social-cooperation differs from individual behavior in hypothalamic and striatal monoamine function: evidence from a laboratory rat model. Behav. Brain Res. *232*, 252–263.
36. Barnett, S.A. (1958). An analysis of social behavior in wild rats. Proc. R. Soc. Lond. *130*, 107–152.
37. Ewer, R.F. (1971). The biology and behavior of a free-living population of black rats (Rattus rattus). Anim. Behav. Monogr. *4*, 127–174.
38. Kashtelyan, V., Lichtenberg, N.T., Chen, M.L., Cheer, J.F., and Roesch, M.R. (2014). Observation of reward delivery to a conspecific modulates dopamine release in ventral striatum. Curr. Biol. *24*, 2564–2568.
39. Steiner, A.P., and Redish, A.D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. Nat. Neurosci. *17*, 995–1002.
40. Stephens, D.W., McLinn, C.M., and Stevens, J.R. (2002). Discounting and reciprocity in an Iterated Prisoner's Dilemma. Science *298*, 2216–2218.
41. Gallistel, C.R., and Gibbon, J. (2000). Time, rate, and conditioning. Psychol. Rev. *107*, 289–344.
42. Brosnan, S.F., Silk, J.B., Henrich, J., Mareno, M.C., Lambeth, S.P., and Schapiro, S.J. (2009). Chimpanzees (Pan troglodytes) do not develop contingent reciprocity in an experimental task. Anim. Cogn. *12*, 587–597.
43. Hillman, K.L., and Bilkey, D.K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. Nat. Neurosci. *15*, 1290–1297.
44. Chang, S.W., Gariépy, J.F., and Platt, M.L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. Nat. Neurosci. *16*, 243–250.
45. Chang, S.W., Winecoff, A.A., and Platt, M.L. (2011). Vicarious reinforcement in rhesus macaques (macaca mulatta). Front. Neurosci. *5*, 27.
46. Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A.J., Schweizer, S., Frith, C.D., and Dalgleish, T. (2009). A key role for similarity in vicarious reward. Science *324*, 900.
47. Dölen, G., Darvishzadeh, A., Huang, K.W., and Malenka, R.C. (2013). Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. Nature *501*, 179–184.
48. Chang, S.W., Barter, J.W., Ebitz, R.B., Watson, K.K., and Platt, M.L. (2012). Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (Macaca mulatta). Proc. Natl. Acad. Sci. USA *109*, 959–964.

Current Biology

Supplemental Information

Prosocial Choice in Rats Depends on Food-Seeking Behavior Displayed by Recipients

Cristina Márquez, Scott M. Rennie, Diana F. Costa, and Marta A. Moita

Supplemental Figures

Figure S1, related to Figure 1

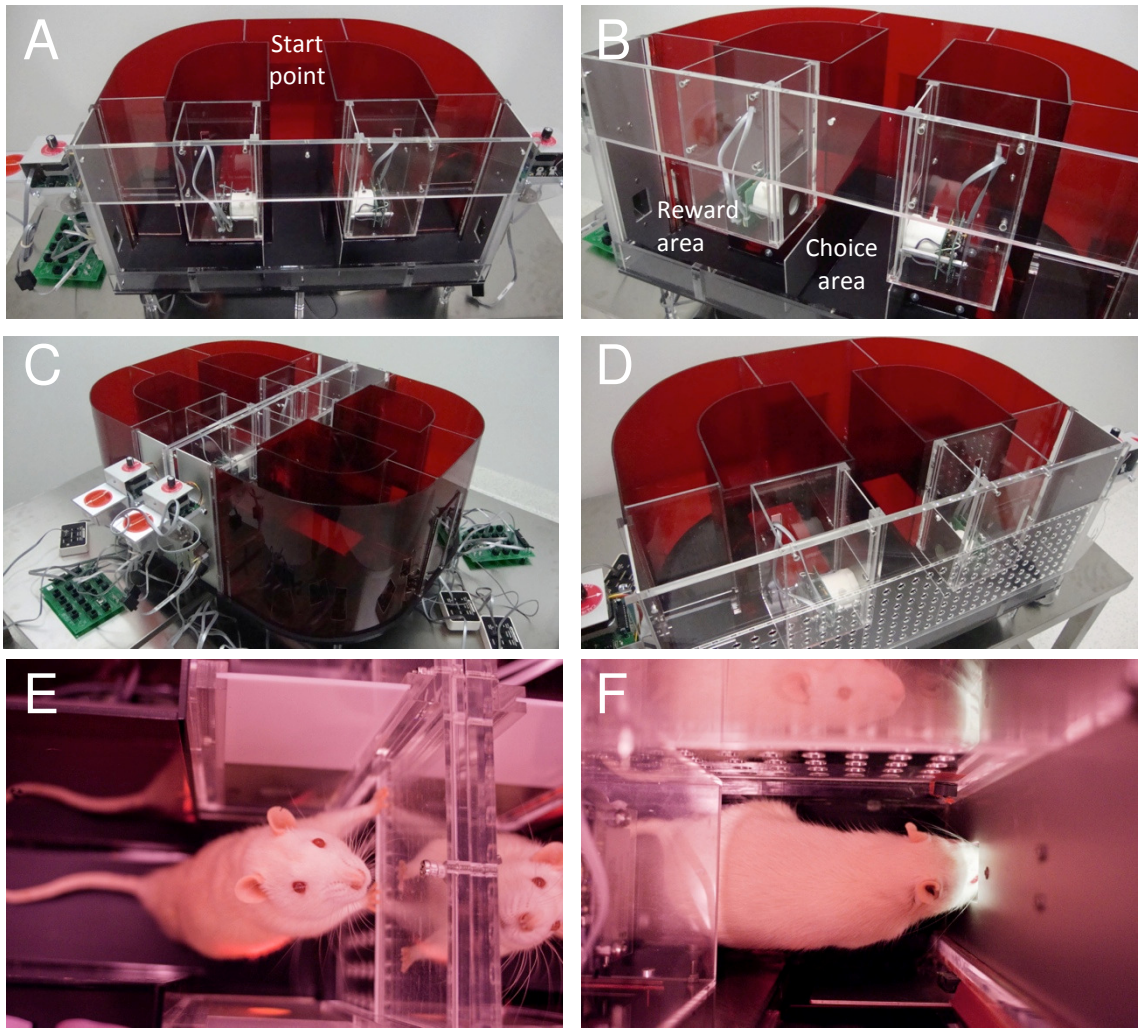


Fig S1: Behavioural apparatus

Our behavioural apparatus consist of two identical fully automated individual T-mazes (**A**-**B**) that can be placed together for the Prosocial Choice Task (**C**).

Each T-maze had a central corridor as starting point (**A**), and two lateral choice arms at the end of which there was a food magazine (reward area, **B**). To gain access to the lateral arms, rats had to poke in a nose port thereby triggering the opening of an automated door. Once in the lateral arm, rats could retrieve food (palatable pellets) and through a small runway go back to the start point to initiate another trial. An automated door placed in the central corridor allowed the control of trial

start by giving access to the choice area (note: this door was removed when the photos were taken, for better visualisation of the entire maze). The fully functional maze can be seen in the Movie S1. Rats were first trained individually to poke in the nose port, retrieve food in the choice arm and go around the maze back to the starting point. Once training was complete the two T-mazes were placed together facing each other (**C**) and the ability for rats to cooperate was tested. Importantly, the wall that separates the two mazes was transparent and perforated (**D**), allowing rats to see, hear, smell and touch each other. Images of two animals performing the prosocial task in our behavioural set up with zoomed views of choice (**E**) and reward areas (**F**) are presented. In **E** two rats rear facing each other in the decision area. The two nose pokes of the recipient rat are visible in the photo. In **F** the recipient rat is retrieving reward (food pellet) while the focal animal is observing from the other maze. The transparent and perforated dividing wall facilitates interaction between the animals.

Figure S2, related to Figure 2.

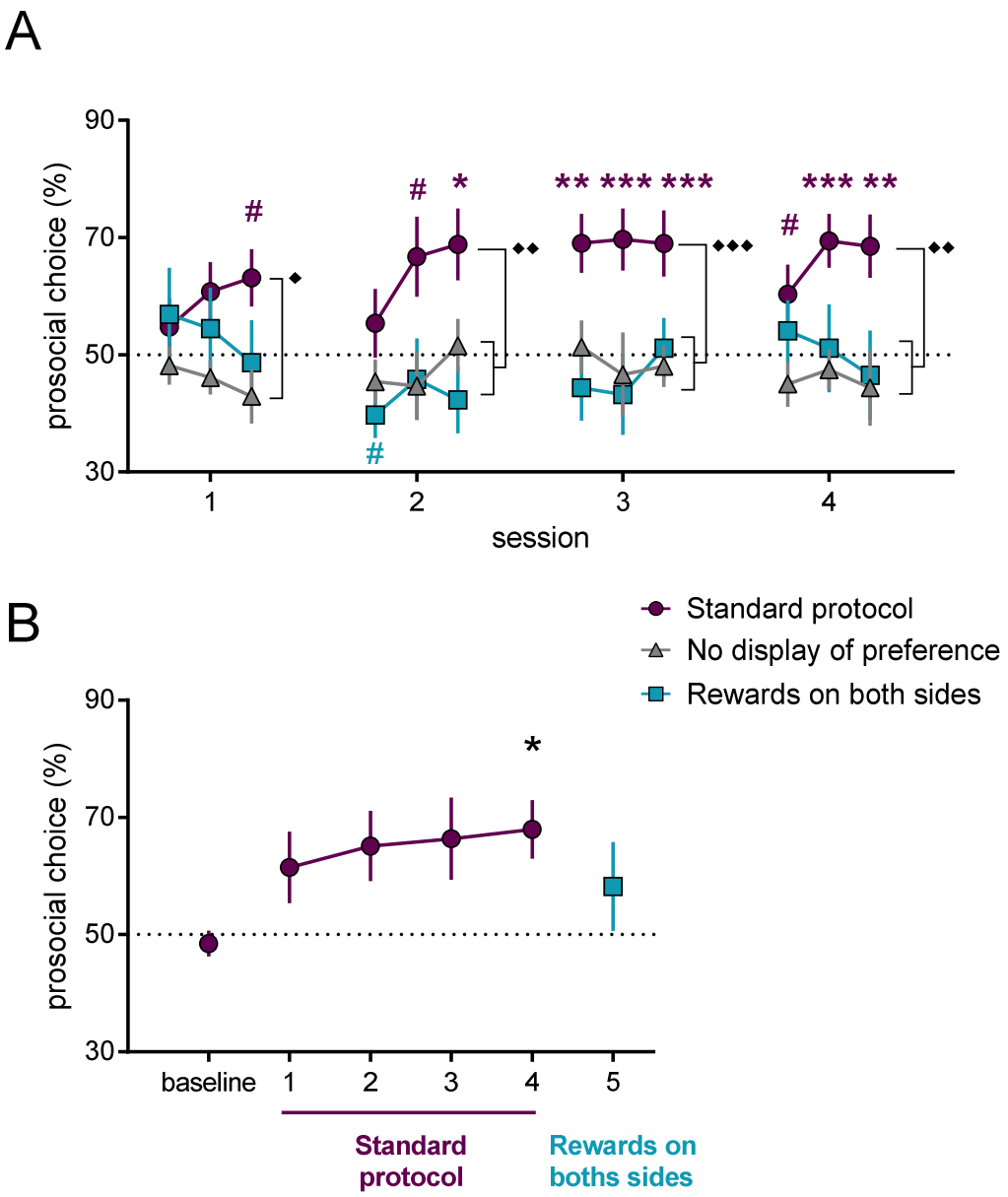


Fig S2: Prosocial Choice in 'standard' protocol emerges within session and is a goal directed action.

(**A**) In order to study the evolution of prosocial choices within each session we divided performance during each testing session in thirds (total number of trials in one session for each dyad divided by three). The percentage of prosocial choices for each third in each experimental protocol was then compared to chance. One sample t-test revealed that differences against chance only emerged in the 'standard' protocol. Regarding the first day of testing no significant differences against chance were observed in the initial phase of testing, but a marginally significant preference towards

prosocial choice emerged at the end of the session only in the 'standard' protocol (see Table S1 for statistical values). Further increases in prosocial choice were observed in the following testing sessions, only in the 'standard' protocol.

Additional statistical analyses were performed in order to compare evolution of choices between experimental protocols within each session. Repeated measures ANOVA with experimental protocol, 'protocol', as between subject factor and each third of the session, 'thirds', as within subject factor were performed independently for each session. In session 1, the 'standard' protocol differed significantly from the 'no display of preference' condition, independently of the moment within the session, and the 'rewards on both sides' protocol showed prosocial choice levels in between these two groups, not reaching significant differences with either protocol (session 1: 'thirds' $F_{(2,68)}$=0.154 p=0.858; 'protocol' $F_{(2,34)}$=3.576 p=0.039; 'thirds' x 'protocol' $F_{(4,68)}$=0.848 p=0.500. Further LSD posthoc analysis revealed significant differences between 'standard' and 'no display of preference' conditions p'=0.011). From the second session of testing onwards, 'standard' protocol choices were significantly higher within the entire session compared to the other two control conditions (session 2: 'thirds' $F_{(2,68)}$=2.391 p=0.099; 'protocol' $F_{(2,34)}$=5.922 p=0.006; 'thirds' x 'protocol' $F_{(4,68)}$=0.491 p=0.491; session 3: 'thirds' $F_{(2,68)}$=0.080 p=0.749; 'protocol' $F_{(2,34)}$=10.064 p=0.00037; 'thirds' x 'protocol' $F_{(4,68)}$=0.482 p=0.804; session 4: 'thirds' $F_{(2,68)}$=0.597 p=0.553; 'protocol' $F_{(2,34)}$=5.484 p=0.009; 'thirds' x 'protocol' $F_{(4,68)}$=1.335 p=0.266). Mean ± SEM are represented**. #** p'<0.10 the colour of the symbol indicates the experimental group of the comparison**, *** p'<0.05 'standard' protocol against chance, ****** p'≤0.01 'standard' protocol against chance, ******* p'≤0.005 'standard' protocol against chance. ◆ between protocols differences p<0.05, ◆◆ between protocols differences p<0.01, ◆◆◆ between protocols differences p<0.005. (**B**) In an independent set of rats, pairs of animals were tested in the 'standard' protocol for four consecutive sessions, in a similar manner as described in the main experiment (n=4). Again, as seen in first experiment using the 'standard' protocol, by the 4th session a preference for the prosocial choice was observed when compared to chance (one sample t-test for the 4th prosocial test session comparing against chance; $t_{(4)}$=3.584 p'=0.046). To test whether rats were sensitive to the contingency between their choice and reward delivered to the recipient, in the fifth session, reward contingencies were changed, such that in this session rewards to the recipients were delivered on both sides (rewards to focal rats remained unchanged, i.e. one pellet on each side). Focal animals were sensitive to this change in experimental design and a drop in the preference for the previous 'prosocial' side was observed. Prosocial choices in this session did not differ significantly from chance (one sample t-test, testing 'rewards on both sides' session against chance $t_{(4)}$=1.081 p'=0.341). These results suggest that

prosocial choice is not a habitual behaviour and that focal animals update their choices depending on the contingency between side chosen and outcome to recipient. Mean ± SEM are represented. * p'<0.05. Corrected p values (p') after sequential Bonferroni correction for multiple comparisons are reported; 2 comparisons were performed.

Table S1: One sample t-test comparing prosocial choice against chance within session (t and p values).

Experimental condition	Session	**1st third**	**2nd third**	**3rd third**
Standard protocol	1	$t_{(14)}$= 0.95 p'=0.358	$t_{(14)}$= 2.12 p'=0.104	$t_{(14)}$= 2.66 p'=0.056 #
	2	$t_{(14)}$= 0.92 p'=0.372	$t_{(14)}$= 2.43 p'=0.058 #	$t_{(14)}$= 3.05 p'=0.026 *
	3	$t_{(14)}$= 3.75 p'=0.006 **	$t_{(14)}$= 3.70 p'=0.005 ***	$t_{(14)}$= 3.34 p'=0.005 ***
	4	$t_{(14)}$= 2.03 p'=0.062	$t_{(14)}$= 4.20 p'=0.003 ***	$t_{(14)}$= 3.42 p'=0.008 **
No display of preference	1	$t_{(10)}$=-0.55 p'=0.592	$t_{(10)}$=-1.33 p'=0.213	$t_{(10)}$=-1.52 p'=0.160
	2	$t_{(10)}$=-1.23 p'=0.249	$t_{(10)}$=-0.92 p'=0.377	$t_{(10)}$= 0.34 p'=0.739
	3	$t_{(10)}$= 0.29 p'=0.775	$t_{(10)}$=-0.46 p'=0.652	$t_{(10)}$=-0.54 p'=0.599
	4	$t_{(10)}$=-1.27 p'=0.232	$t_{(10)}$=-0.73 p'=0.485	$t_{(10)}$=-0.86 p'=0.410
Rewards on both sides	1	$t_{(10)}$= 0.86 p'=0.410	$t_{(10)}$= 0.64 p'=0.536	$t_{(10)}$=-0.18 p'=0.860
	2	$t_{(10)}$=-2.60 p'=0.079 #	$t_{(10)}$=-0.60 p'=0.561	$t_{(10)}$=-1.34 p'=0.209
	3	$t_{(10)}$=-1.01 p'=0.337	$t_{(10)}$=-0.98 p'=0.349	$t_{(10)}$= 0.22 p'=0.829
	4	$t_{(10)}$= 0.78 p'=0.453	$t_{(10)}$= 0.14 p'=0.885	$t_{(10)}$=-0.46 p'=0.657

p'<0.10**, *** p'<0.05, **** p'≤0.01, ***** p'≤0.005. Corrected p values (p') after sequential Bonferroni correction for multiple comparisons are reported; 3 comparisons were performed in each testing session.

Figure S3: related to Figure 3

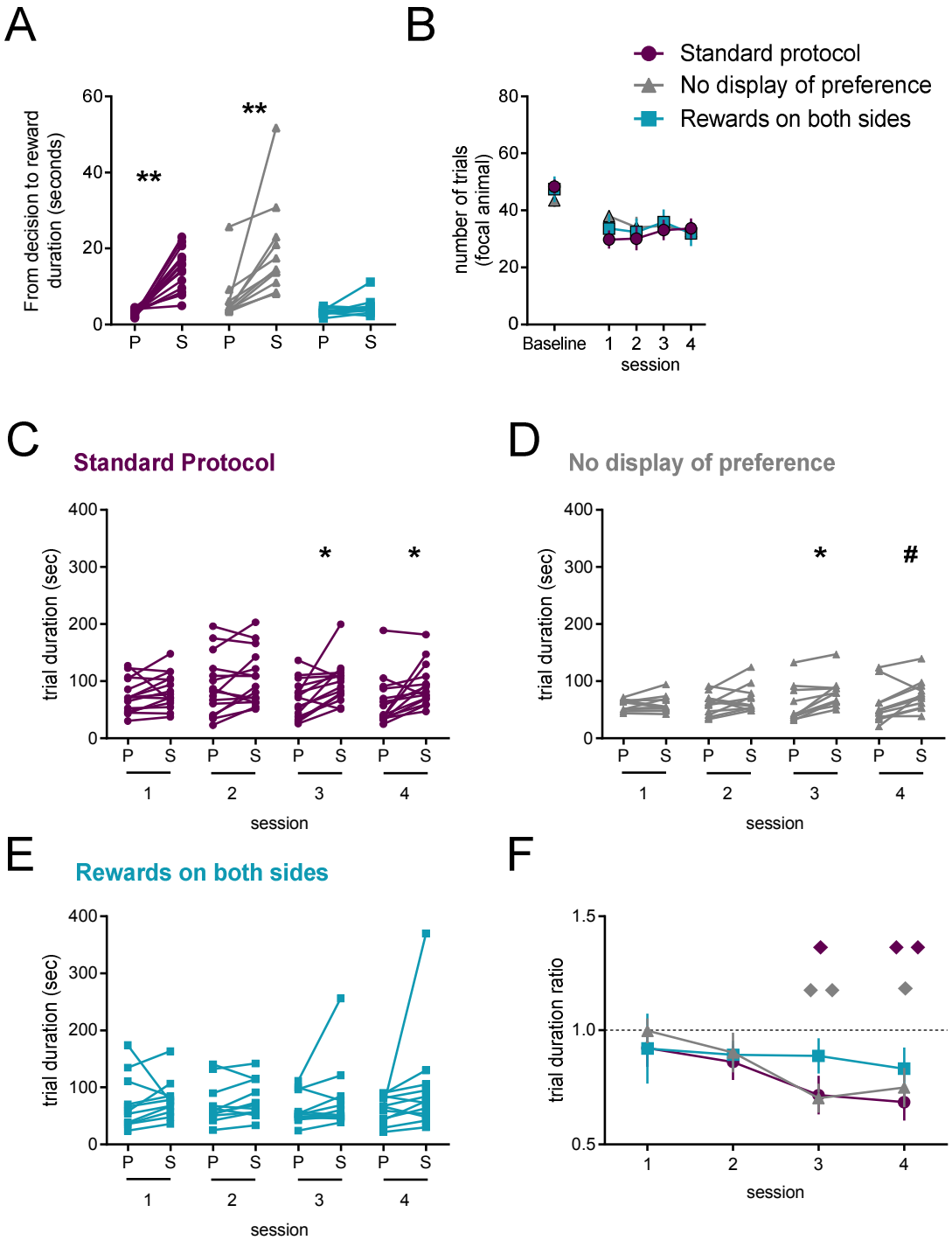


Fig S3: Differences in number of trials and trial duration between prosocial and selfish trials.

Time elapsed between door opening (upon decision by the focal rat) and entry of both rats in the reward area is shown for prosocial (P) and selfish (S) trials in the three experimental groups (**A**). Rats took longer to enter the selfish arm, only in the experimental conditions where the recipient would not receive pellets in that arm (non-parametric test for paired samples Wilcoxon signed ranks test

were performed between prosocial and selfish trials within each experimental protocol; Z=-3408 p'=0.002 for 'standard' protocol, Z=-2.934 p'=0.007 for 'no display of preference' protocol, Z=-1.334 p'=0.182 for 'Rewards in both sides' protocol). Total number of trials performed during baseline or Prosocial Choice Test was similar among experimental protocols (**B**), indicating that the number of rewards received by focal animals (one per trial) was comparable in the three conditions (Baseline: one way ANOVA testing the effect of experimental protocol, $F_{(2,34)}$=0.650 p=0.528; Prosocial Choice testing sessions: repeated measures ANOVA with 'experimental protocol' as between subjects factor and 'session' as within-subjects factor, revealed no significant effect of 'experimental protocol' ($F_{(2,34)}$=0.203 p=0.817), 'session' ($F_{(3,102)}$=1.229 p=0.303), nor the interaction 'protocol' by 'session' ($F_{(6,102)}$=1.815 p=0.103)). Total trial duration over sessions for the two trial types is plotted for the 'standard' protocol (**C**), 'no display of preference' (**D**) and 'Rewards in both sides' (**E**) conditions. Non-parametric test for paired samples Wilcoxon signed ranks test was performed revealing that selfish trials were longer than prosocial ones in 'standard' and 'no display of preference' conditions, and only significant from session three onwards (see supplementary Table S2 for statistical values) . When the ratio prosocial/selfish trials duration was calculated (**F**), no differences between experimental conditions were observed (repeated measures ANOVA with experimental protocol, 'protocol' as between subjects factor and testing session, 'session' as within subject factor, revealed a significant effect of 'session' ($F_{(3,102)}$=3.512 p=0.018), no significant effect of 'protocol' ($F_{(2,34)}$=1.154 p=0.327) nor interaction 'protocol' x 'session' ($F_{(6,102)}$=0.524 p=0.789)). Moreover, to assess in which session a significant relative difference between the duration of prosocial and selfish trials emerged, we performed a one-sample t-test against 1 (a ratio of 1 means that duration of prosocial trials equals duration of selfish ones) for each session. We found that differences in trial duration emerged only on the third day of prosocial testing and on those groups where reward for the recipient was available only in one arm (one-sample t-test against 1: standard' protocol: $t_{(14)}$=-0.95 p'=0.359 for session 1; $t_{(14)}$=-1.75 p'=0.199 for session 2; $t_{(14)}$=-3.36 p'=0.014 for session 3; $t_{(14)}$=-3.87 p'=0.007 for session 4; 'no display of preference' protocol: $t_{(10)}$=-0.04 p'=0.966 for session 1; $t_{(10)}$=-1.11 p'=0.293 for session 2; $t_{(10)}$=-4.69 p'=0.003 for session 3; $t_{(10)}$=-2.90 p'=0.047 for session 4; 'rewards on both sides' protocol: $t_{(10)}$=-0.52 p'=0.618 for session 1; $t_{(10)}$=-1.75 p'=0.111 for session 2; $t_{(10)}$=-1.43 p'=0.182 for session 3; $t_{(10)}$=-1.79 p'=0.103 for session 4). n= 15 for 'standard' ' protocol and n=11 for 'no display of preference' and 'rewards on both sides' protocols. Mean ± SEM are represented. **p'<0.01, *** p'≤0.005. ◆ p'<0.05 and ◆◆ p<0.01 after one-sample t-test against 1 as value of reference, the colour of the symbol indicates the experimental protocol of comparison. Corrected p values (p') after sequential Bonferroni correction for multiple comparisons are reported; 3 comparisons were performed in A and 4 comparison in C-F.

Table S2: Comparisons between selfish and prosocial total trial duration over sessions. Non-parametric tests for paired samples: t and p values obtained.

Experimental condition	Session	Statistics
Standard protocol	1	*Z* = -1.761 p'=0.156
	2	*Z* = -1.429 p'=0.155
	3	*Z* = -2.612 p'=0.027 *
	4	*Z* = -2.839 p'=0.018 *
No display of preference	1	*Z* = -0.711 p'=0.477
	2	*Z* = -1.334 p'=0.182
	3	*Z* = -2.756 p'=0.023 *
	4	*Z* = -2.133 p'=0.098 #
Rewards on both sides	1	*Z* = -1.336 p'=0.182
	2	*Z* = -1.511 p'=0.131
	3	*Z* = -1.245 p'=0.213
	4	*Z* = -1.867 p'=0.247

p'<0.10**, *** p'<0.05. Corrected p values (p') after sequential Bonferroni correction for multiple comparisons are reported; 4 comparisons were performed in each experimental protocol.

Figure S4, related to Figure 4:

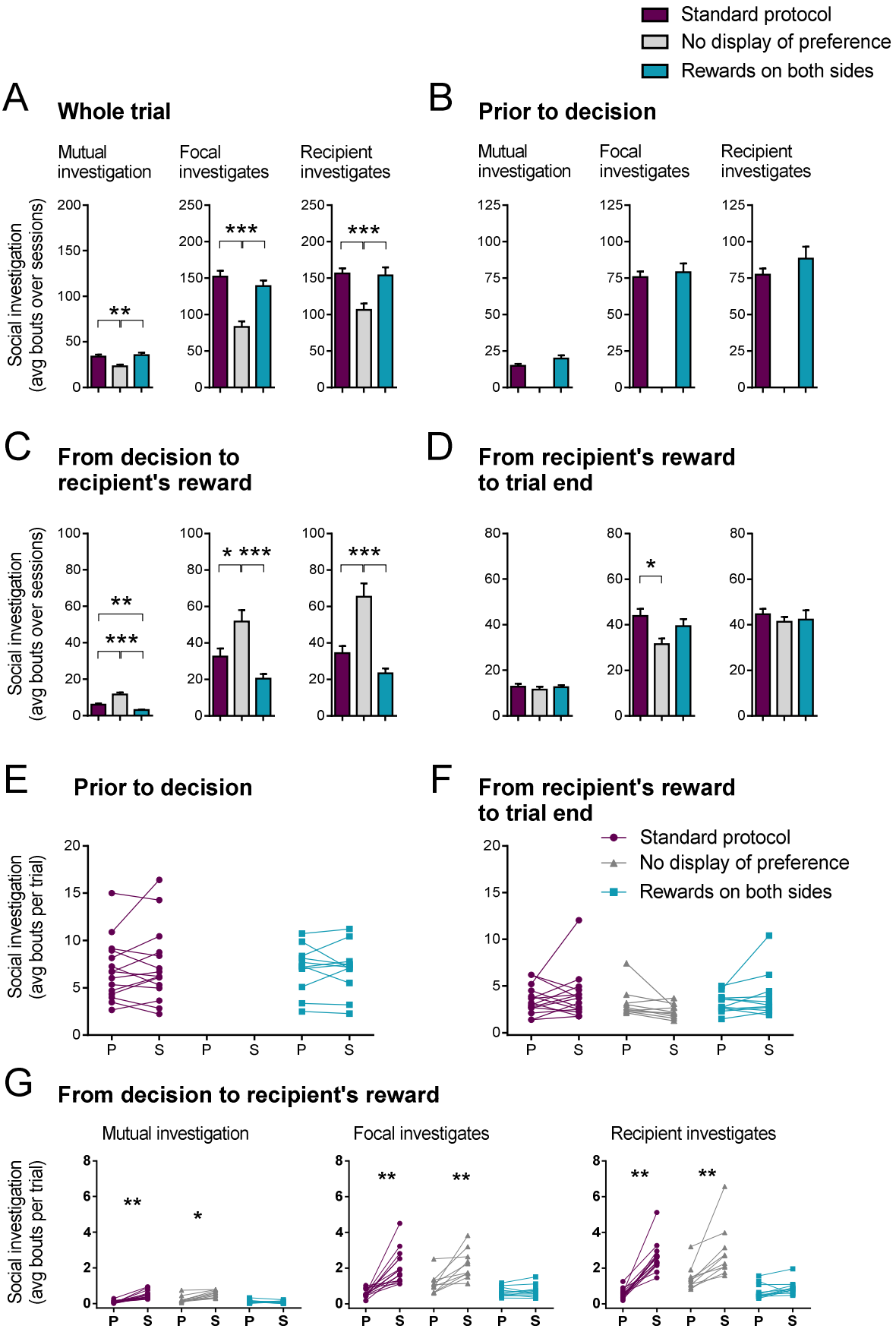


Fig S4: Social investigation sub-types show similar pattern as total social investigation.

Average number of bouts per session of the different types of social investigation (mutual investigation, focal investigates recipient and recipient investigates focal) are shown for whole trial (**A**), and for different segments of a trial: prior to decision (**B**), from decision until the recipient animal retrieved the reward (or in case of selfish trials, when it entered the unrewarded arm) (**C**) and from recipient's reward until both animals were back in the central corridor ending the trial (**D**). Measurements of the three different types of social interaction revealed a similar pattern to that observed when the sum of all types of behaviours was used (see main **Fig 4**). Social investigation measured in the Whole trial (**A**) was decreased in the 'no display of preference' protocol in all types of social investigation measured (One way ANOVA; Mutual investigation $F_{(2,33)}$=6.88 p=0.003; Focal investigates recipient $F_{(2,33)}$=19.11 p<0.0001; Recipient investigates focal $F_{(2,33)}$=9.46 p=0.001). Further LSD posthoc analysis for the Whole trial revealed that social investigation was decreased in the 'no display of preference' protocol compared to the 'standard' protocol (Mutual investigation p'=0.007; Focal investigates recipient p'<0.0001; Recipient investigates focal p'=0.0005) and 'rewards on both sides' protocol (Mutual investigation p'=0.005; Focal investigates recipient p'=0.0001; Recipient investigates focal p'=0.003). No differences were observed between 'standard' and 'rewards on both sides' protocols when social investigations in the Whole trial were studied (LSD posthoc: Mutual investigation p'=0.593; Focal investigates recipient p'=0.245; Recipient investigates focal p'=0.829). All types of social investigation prior to decision (**B**) were similar in the two groups that were allowed to interact, as observed with the summed number of interactions (see main **Fig 4**) (t-test; Mutual investigation $t_{(24)}$=-1.93 p=0.066; Focal investigates recipient $t_{(24)}$=-0.50 p=0.622; Recipient investigates focal $t_{(24)}$=0.20 p=0.204). From the moment of focal's decision to the moment that both animals were in the reward areas (**C**), an increase in the amount of all types of social interaction was observed in the 'no display of preference' protocol, maybe as a compensation for the lack of social interaction prior to decision (One way ANOVA; Mutual investigation $F_{(2,33)}$=29.18 p<0.0001; Focal investigates recipient $F_{(2,33)}$=10.46 p=0.0003; Recipient investigates focal $F_{(2,33)}$=18.45 p<0.0001). Further LSD posthoc analysis revealed that animals on the standard' protocol displayed higher levels of mutual investigation than those from 'rewards on both sides' protocol (p'=0.008) and that social investigation was increased in the 'no display of preference' protocol compared to the 'standard' protocol (Mutual investigation p'<0.0001; Focal investigates recipient p'=0.011; Recipient investigates focal p'=0.0001) and 'rewards on both sides' protocol (Mutual investigation p'<0.0001; Focal investigates recipient p'=0.0002; Recipient investigates focal p'<0.0001). Social investigation from recipient's reward to trial end (**D**) was similar in all groups as observed with the summed number of interactions (see main **Fig 4**), except for the case of Focal investigates Recipient

were ‘standard’ protocol investigated more than the ‘no display of preference’ protocol (One way ANOVA; Mutual investigation $F_{(2,33)}$=0.27 p=0.768; Focal investigates recipient $F_{(2,33)}$=4.18 p=0.024; Recipient investigates focal $F_{(2,33)}$=0.34 p=0.711; further LSD posthoc revealed higher levels of social investigation from the focal animal in the ‘standard’ protocol compared to the ‘no display preference’ condition p’=0.02). No differences in the average number of social investigation bouts depending on trial type were observed prior to decision (**E**) nor from the moment of the recipient’s reward to trial end (**F**). Non-parametric tests for paired samples Wilcoxon signed ranks test were run to compare the amount of social investigation observed in prosocial (P) and selfish (S) trials for each protocol. (**E**) Prior to decision: ‘standard’ protocol *Z*=-0.568 p’=0.570; ‘rewards on both sides’ *Z*=-0.089 p’=0.929. (**F**) From recipient’s reward to trial end: ‘standard’ protocol Z=-0.454 p’=0.650; ‘no display of preference’ *Z*=-1.912 p’=0.280; ‘rewards on both sides’ Z=-1.156 p’=0.248. Measurements of the different sub-types of social investigation from focal’s decision until both animals were in the reward area reflected the same differences as observed in the total social investigation (**G**; see also main **Fig 4**). Mutual investigation, investigations made by focal rat and investigations performed by recipient animal were higher in selfish trials in the ‘standard’ and ‘no display of preference’ protocols, where recipient animals were not being rewarded. Non-parametric tests for paired samples Wilcoxon signed ranks test were run for the amount of each sub-type of social investigation observed in prosocial (P) and selfish (S) trials for each protocol. Mutual investigation: ‘standard’ protocol Z=-3.41 p’=0.002; ‘no display of preference’ Z=-2.80 p’=0.015; ‘rewards on both sides’ Z=-0.80 p’=0.423. Focal investigates recipient: ‘standard’ protocol Z=-3.41 p’=0.001; ‘no display of preference’ Z=-2.80 p’=0.010 ; ‘rewards on both sides’ Z=-0.80 p’=0.424. Recipient investigates focal: standard’ protocol Z=-3.41 p’=0.0006; ‘no display of preference’ Z=-2.80 p’=0.005 ; ‘rewards on both sides’ Z=-1.96 p’=0.150. *p’<0.05, **p’<0.01 Mean ± SEM (A-D) and individual values (E-G) are represented. *p’<0.05, **p’<0.01, *** p’<0.005. Corrected p values (p’) after sequential Bonferroni correction for multiple comparisons are reported; 3 comparisons were performed in each graph (except for data in panel **B**).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Animals

74 adult male Sprague-Dawley rats (Charles-River, France) were used in the experiments (375-425 g body weight at the beginning of the experiment). Upon arrival, rats were pair-housed and maintained with *ad libitum* access to food and water under a reversed light cycle (12 hours dark/light cycle; lights off at 10 AM) in controlled temperature conditions. Animals were left undisturbed in their home-cages for three weeks, allowing rats to habituate to our Vivarium Facility and routines, and to reverse their circadian rhythm. After this period, animals were handled during three days and habituated to the food pellets used in our task, in the home-cage. Experiments were performed during the second phase of the dark cycle, i.e. at least 3 hours after the lights were off. Animals were not food restricted during testing. Experimental procedures follow the Portuguese Guidelines, which comply with the European Directive 86/6097EEC of the European Council.

Behavioural apparatus

The behavioural apparatus consisted of two identical fully automated individual T-mazes (**Fig S1A-B**) that could be placed together for the Prosocial Choice Task (**Fig S1C**). Custom made acrylic mazes (70 cm x 40 cm, 30 cm high each individual maze) (Gravoplot, Sintra, Portugal) where automatically controlled using Graphic State 3.03 software and the Habitest interface (Coulbourn Instruments, Allentown, PA, USA). Custom-made automatic doors (Champalimaud Scientific Hardware Platform, Lisbon, Portugal; Mobiara R&D Design, Lisbon, Portugal; WGT-Elektronik, Kolsass, Autria) triggered by infrared beams (Champalimaud Scientific Hardware Platform, Lisbon, Portugal) controlled the position of the animals in the mazes. Each T-maze had a central corridor as starting point (**Fig S1A**), and two lateral choice arms at the end of which there was a food magazine. In the central corridor an automated door determined the beginning of the trial by giving access to the decision area, where the entrances to the lateral arms were located. To gain access to the lateral arms, rats had to poke in a light-cued nose port (Coulbourn Instruments, Allentown, PA, USA) placed above an automated door triggering its opening. Once in the lateral choice

arm, rats could retrieve food (45 mg pellet, reference F0021, BioServ, Frenchtown, NJ, USA), triggering the opening of the door that gave access to a small runway leading to the starting point at the centre arm, thus initiating another trial.

Rats were first trained individually to poke in the nose port, retrieve food in the choice arm and go around the maze back to the starting point. Once training was complete the two T-mazes were placed facing each other (**Fig S1C**) and the ability of rats to cooperate was tested. Importantly, the wall that separates the two mazes was transparent and perforated (**Fig S1D**), allowing rats to see, hear, smell and touch each other.

Individual training in standard protocol

All animals were first habituated to the individual T-mazes, in two 15 min free-exploration sessions, where all doors were open and nose pokes and infrared beams inactive. After the second session, pellets were delivered in the food magazines and free consumption was allowed during 5 min. Next, rats were habituated to the opening and closing of the automated doors, by having these open and close independently of the rats' behaviour. Once habituation sessions finished, all animals were shaped to rear to poke in the nose port in order to open the door that gave access to the food magasin. Rats were allowed to nose poke and explore both arms, which were always rewarded with several pellets. In order to increase the motivation to explore the maze and search for food, rats were slightly food-restricted in the following manner: number of food pellets in the home-cage was restricted to a 90% of the baseline intake, during the period from lights off to the training. Rats were allowed to eat *ad libitum* for the rest of the cycle. This minimal food restriction was only applied at the beginning of individual training (two days maximum) and no decrease in body weight was observed. Once animals performed few trials in this schema, within each dyad, rats were randomly assigned to be the helper (henceforth the focal rat) or the recipient of help. From that moment onwards focal and recipient rats received different kinds of individual training (see bellow). Focal and recipient roles were fixed throughout the entire experiment.

Individual training for focal rats

Focal animals were trained in the individual mazes during several sessions to perform one poke in the nose port, retrieve one food-pellet in the choice arm and go around the maze

back to the starting point. During training, focal rats received the same amount of food in both arms, until they were equally likely to visit either arm (See **Fig 1F**). Rats tend to perform alternations, and no side preference was observed at the end of the training. Focal animals were trained for at least 5 sessions. In the last two sessions of individual training, a delay in the opening of the central door was included, such that focal animals would have to wait a pseudo-random period (from 2 to 20 seconds) to have access to the choice area. The rationale for this delay in starting the trial was to habituate focal animals to wait in the start area, as they might have to do in the Prosocial Choice Task, when they would have to coordinate with the recipient rat (a trial only started when both rats were in the centre arm and the nose ports of focal rats were only active after the recipient made two nose pokes).

Individual training for recipient rats

The aim of individual training for recipients was to (i) show a clear preference for one side of the maze and (ii) display a vigorous food-seeking behaviour (in our Prosocial Choice Task, nose poking). To this end, recipients were rewarded in one arm only (randomly assigned to be the right or left arm) cued with a light whose nose port was active. In this way, recipients would learn to go to only one side of the maze where they could retrieve one food-pellet (**Fig 1E**, left panel). Nose poke training started with a fix ratio one schedule (FR1: one nose poke in the assigned arm would open the door in that same arm, giving access to reward), and increased over days until FR5 was performed correctly (being the criteria to perform trains of nose pokes and be selective to the active nose poke). During this phase of training recipient animals were minimally food-restricted (90% of baseline food intake, during the hours prior to training). In the last three days of individual training, recipients had to nose poke under a variable ratio five schedule (VR5: an average of 5 nose pokes were needed to open the door of the rewarded arm). In the last two sessions, a delay in the opening of the central door was included, so recipient animals would have to wait a pseudo-random period (from 2 to 20 seconds) to have access to the choice area, in order to habituate them to waiting periods for the focal animal in the Prosocial Choice Task. Moreover, during these last two sessions of training, recipient rats were forced to visit the unrewarded arm in 10 and 20% of the trials. In this manner, recipient rats would learn that even if no pellets were given in the unrewarded arm, they would have to go into the lateral arm and back to the start point to initiate the next trial. Finally, recipients were briefly re-trained immediately

before each session of the prosocial choice task, to avoid extinction of food-seeking behaviour.

Prosocial Choice Task in standard protocol

During the Prosocial Choice Task pairs of cage-mate rats (n=15) were tested in the double T-maze (one per rat) described above, where a centre arm gave access to two food-baited arms gated by automated doors. These doors were controlled by nose ports placed above them, such that when a rat would poke the nose port the door underneath would open (see **Fig 1** and **S1**). Poking the nose port corresponded to food-seeking behaviour in our task. During testing, although both rats had access to the nose ports of their corresponding mazes, only the ones of the focal were active and these controlled the doors of both mazes. In this manner, the recipient rat displayed food-seeking behaviour while the focal controlled the recipient's access to the food-baited arms. A trial started when both focal and recipient rats were in the center corridor. This would trigger opening of the centre doors giving access to the decision areas. At this point the recipient rat had to perform a minimum of two nose pokes (food-seeking behaviour) after which focal's nose ports would become active. Then, the focal rat could choose which nose port to poke. Prosocial choice corresponded to choosing the side of maze that provided access to food to itself and the recipient, whereas choosing the side of the maze that provided food to itself and no food to the recipient corresponded to a selfish choice. Hence, prosocial and selfish choices provided the same amount of food to the focal rats. Importantly, in both choices (prosocial and selfish) focal and recipient rats went to the same side of the maze, so that choice was not affected by a desire for proximity to the recipient of help. Reward to the focal animal was available in its food magazine immediately after decision was made. However, the recipient rat would only receive its pellet once both animals were in the lateral arm, ensuring that information about the recipient getting or not getting reward was available for the focal animal. After entering the lateral arms, both animals would have to go back to the start point in order to reinitiate a new trial, synchronizing the presence of the pairs in the decision areas. Four daily sessions of 40 minutes were performed.

Dissecting motivations of prosocial behaviour: sensitivity to the display of food-seeking behaviour

A modification of the 'standard' protocol was performed in order to evaluate the role of the display of food-seeking behaviour of the recipient rat in the Prosocial Choice Task (n=11). Focal's individual training would be similar as in the 'standard' protocol, but training of recipient rats and Prosocial Choice Task was modified. During individual training in this condition, recipient rats did not need to nose poke to gain access to the baited-arms. Once in the central corridor, one of the lateral doors would automatically open and recipient rats could only visit the open arm (See **Fig 1E**, middle panel). During Prosocial Choice Task, recipients were kept behind a door (transparent and perforated), away of the decision area, from trial start until the focal made its choice. In this way, no display of food-seeking behaviour (nose pokes or other possible types of seeking behaviour) were possible before the focal would nose poke in one of the arms. Once the decision was made, lateral doors of the choice-arm would open for both mazes and the central door would also open for the recipient. In the same way as in the 'standard' protocol, recipients in this condition would only be rewarded in one arm, which was counterbalance over pairs of animals, but kept constant over the experiment (**Fig 2C**, middle panel).

Dissecting motivations of prosocial behaviour: sensitivity to reward delivery

A modification of the 'standard' protocol was performed in order to evaluate the role of reward delivery to the recipient rat in the Prosocial Choice Task (n=11). Focal's individual training and Prosocial Choice Task would be similar as in the 'standard' protocol, but training of recipient rats was modified. During the last three days of individual training, where VR5 was introduced, still only one nose port was active, but now poking the active poke randomly opened either one of the two lateral arms. Both arms would be always rewarded. In this way, during the Prosocial Choice Task, recipients would continue displaying vigorous side-specific nose pokes, but would be rewarded always, independently of the choice of the focal rats (See **Fig 1E** and **2A**, right panels).

Data extraction from Mazes Interfaces

Data from the positions and behaviour of the animals in the automated mazes was extracted from Graphic State 3.03 Software (Coulbourn Instruments, Allentown, PA, USA) and parsed using Matlab R2010b.

Video analysis of social behaviours

Detailed video analysis of interacting pairs of animals in the maze was conducted by an experimenter blind to the treatment conditions, and assisted by a custom-made computer programme that recorded frequency and duration of pre-defined behaviours. Number of times focal and recipient rats investigated each other, number of times the focal unilaterally investigated the recipient and viceversa were quantified. We used number of bouts for the present analysis, as the duration of the social interactions of the pair was very short (average duration of the bouts of social interactions was 0.97 ± 0.017 seconds, with a minimum value of 0.59 and a maximum of 1.31 seconds). Moreover, total duration of social interactions was highly correlated with the number of bouts (Pearson's correlation, $r=0.875$, $p<0.0001$).

Finally, video analysis data was aligned to the events performed in the maze extracted from Graphic State 3.03 Software (Coulbourn Instruments, Allentown, PA, USA). Alignment of the maze events with the social interaction data was not successful in the third session of one animal from the 'no display of preference' group. We thus decided to exclude this animal from all analysis related to social interactions (n=10).

Statistical Analysis

The SPSS 13.0 (SPSS, Chicago, IL, USA) statistical package was used for the statistical analyses. The normality and homogeneity of variance of the data were tested, and the appropriate statistics were used as required.

Choice preference: (i) To test for deviations from chance, 50% choice, one sample t-tests were used. This analysis was performed for 1) individual training sessions, 2) each session of prosocial behaviour testing and 3) for first, middle and last third within each test session. (ii) Changes in prosocial choice were confirmed by comparing the focals' baseline choices to their choices in first test session using a paired samples t-test, being baseline the average of focal choices during the last two days of individual training. (iii) In order to evaluate changes

across sessions in the choices of focal animals from the 'standard' protocol, a repeated measures ANOVA with 'session' as within-subject factor was performed. (iv) Two-way ANOVA, followed by LSD posthoc tests, was performed to compare prosocial choices across the three different protocols over the course of the four testing sessions. (v) To assess the side preference of each individual rat, an exact test was used computing the probability that the observed proportion of prosocial and selfish choices (taken from all trials of all sessions) could arise from chance. For p values $p<0.05$ rats were considered prosocial (in cases where prosocial choices were more frequent) or selfish (where selfish choices were more frequent), for $p>0.05$ rats were considered unbiased.

Nose pokes: Non-parametric Mann-Whitney test was used to evaluate differences between the number of nose pokes displayed by recipient rats.

Number of trials: Two-way ANOVA with 'experimental protocol' as between subjects factor and 'testing session' as within-subjects factor was performed to compare number of trials performed (and thus, number of pellets eaten by focal animals).

Trial duration: (i) Non-parametric paired tests (Wilcoxon signed ranks test) were performed to study differences in the duration of trials between prosocial and selfish trials within each experimental protocol. (ii) One-way ANOVA was used to evaluate differences in the trial duration ratio across protocols. (iii) Pearson correlations were performed between trial duration ratio and focal choices. (iv) one sample t-test against reference value of 1 was used to evaluate the emergence of a significant relative difference in the durations of prosocial and selfish trials for each testing session. (v) the same analysis as in ii-iv were performed using absolute difference (length of selfish trial – length of prosocial trial).

Social investigation: (i) One-way ANOVA tested the effect of experimental protocols on social investigation, followed by LSD posthoc tests when needed. ii) Non-parametric paired tests (Wilcoxon signed ranks test) were performed to study differences between prosocial and selfish trials within each experimental protocol.

Sequential Bonferroni correction was performed for all pairwise comparisons to correct for multiple comparisons and corrected p values (p') are reported. Statistical significance was set at adjusted p value $p'<0.05$. Uncorrected p-values were first rank-ordered by significance (from smallest to highest p value) and then corrected using the following algorithm (for m = number of hypothesis tested, i = hypothesis number, j = ranked position of p value):

Brains.